X-RAY DIFFRACTION STUDIES OF GALACTOSE DERIVATIVES AND POLYMERS

by

John W. Campbell

THESIS

presented for the Degree of Doctor of Philosophy

of the

UNIVERSITY OF EDINBURGH

December, 1969

EX BIBL UNIV. EDINBURGEN
CHEMISTRY LIBRARY

TO MY PARENTS

SUMMARY

Iota-carrageenan is a sulphated seaweed polysaccharide with a repeating structure of the type $(-B-A-)_n$ where B is a 1, 3 linked β-D-galactose-4-sulphate residue and A is a 1, 4 linked 3, 6 anhydro-α-D-galactose-2- sulphate residue. In the presence of certain cations iota-carrageenan forms strong gels which may be drawn into oriented fibres. X-ray diffraction photographs of these fibres and of fibres of the closely related polysaccharide kappa-carragenan indicate that iota-carrageenan forms a double helix with a pitch of 26.0 Å and with three disaccharide residues per turn of helix. The second chain is parallel to the first and displaced from it by half a helix repeat distance parallel to the helix axis giving an observed fibre axis repeat distance of 13.0 Å. The double helices are packed in a hexagonal array. Model building calculations confirmed that a right handed double helix of the type described above was sterically feasible and indicated a possible conformation for the polysaccharide chains. Intensity calculations, packing considerations and the presence of a hydrogen bond for which there was experimental evidence indicated that the proposed model was reasonable. Cylindrically averaged Patterson functions were calculated in order to examine the packing arrangement in more detail and to try to locate the cations. The results from these were however rather inconclusive.

The crystal structure of methyl-3, 6-anhydro-α-D-galactoside, angal, was determined in order to obtain an accurate set of coordinates for 3, 6-anhydrogalactose for model-building calculations

/on ...

on iota-carrageenan. Angal crystallised in space group $P2_1 2_1 2_1$ with a= 9.46 Å, b = 12.05 Å and c = 6.93 Å. After an unsuccessful attempt to solve the structure in projection, three dimensional x-ray diffraction intensity data (746 relfections) were collected on film. Application of direct methods, using the symbolic addition procedure and the tangent formula, led to the solution of the structure. The structure was refined using a full matrix least squares procedure. A second set of intensity data (1974 reflections) was collected on a computer controlled four circle diffractometer and the structure was re-refined. The final R factor was 0.097 and the standard deviations were 0.004 Å in the bond lengths and 0.2 degrees in the bond angles.

Part of the work described in Part II of this thesis has been published in collaboration with Dr. N.S. Anderson, Dr. M.M. Harding, Dr. D.A. Rees and Dr. J.W.B. Samuel. A copy of the paper is included at the end of the thesis.

CONTENTS

SECTION VII: OTHER METHODS TRIED.

PART II: THE STRUCTURE OF IOTA CARRAGEENAN.

PART I

THE CRYSTAL STRUCTURE OF METHYL-3,6-ANHYDRO-α-D-GALACTOSIDE (ANGAL)

SECTION I: THE CRYSTAL DATA

1.1, INTRODUCTION

Model-building methods, based on a knowledge of the geometry of the constituent monosaccharide residues, play an important part in determining polysaccharide chain conformations. 3, 6-Anhydrogalactose is one of the residues which form the backbone of the polysaccharide iota-carrageenan (see Part II of thesis) and, in order to obtain an accurate set of coordinates for this residue, the crystal structure determination of methyl-3, 6-anhydro-α-D-galactoside, henceforth referred to as angal, was undertaken. A set of coordinates was estimated for angal before the results of the crystal structure determination were available and this was used in model-building calculations on iota-carrageenan and also for various methods used in attempting to solve the angal structure. This model proved to be a close approximation to the experimentally derived structure.

Angal crystallised in space group $P2_1\ 2_1\ 2_1$ and an attempt was made to solve the structure in projection. Methods involving a systematic exploration of possible molecular orientations and positions in the unit cell were tried, as this type of approach would probably have to be used in the polysaccharide work where only a limited quantity of data was available. The projection work did not lead to a solution of the structure and a three dimensional analysis, based on film data, was undertaken. Visual inspection of a three dimensional

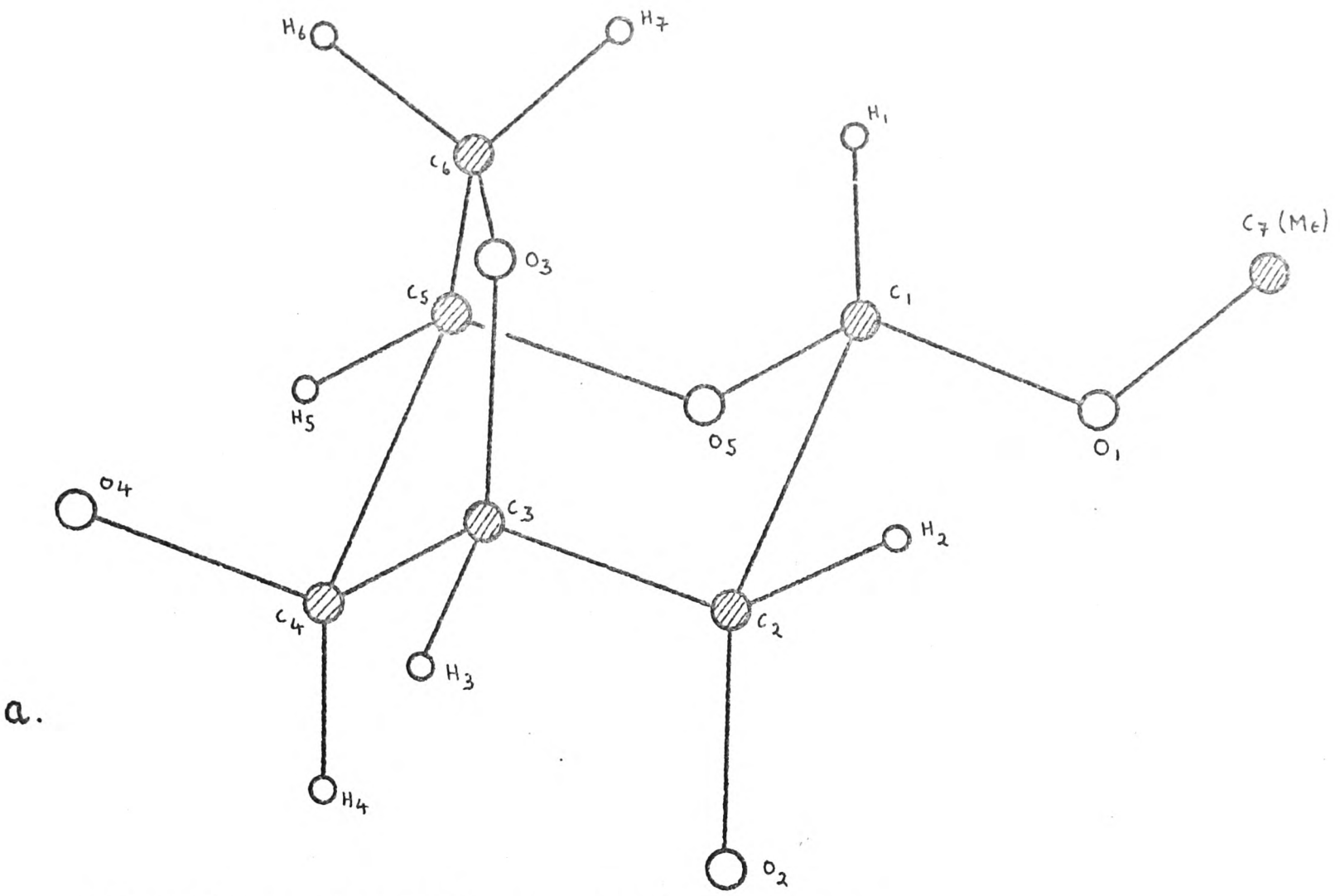
/Patterson ...

Patterson function in the region near to the origin gave useful information about the orientation of the molecule, though various rotation functions, designed to explore orientations systematically, were not helpful. The application of direct methods led to the solution of the structure which was refined using a full matrix least squares procedure. As some of the intensity measurements were unsatisfactory and as the quantity of data was rather small, a second set of intensity data was collected on a Hilger and Watts four circle diffractometer and a second refinement carried out.

In sections I to VI of this thesis on the solution, refinement and discussion of the angal structure, only a brief summary of the projection work is included. This work and other methods which did not lead directly to the solution of the structure are described in Section VII.

Angal (Fig. 1a) is also of interest in the effect that the strain within the molecule has on its chemical reactions. The reactions of angal and related compounds have been studied by Haworth et al. (1). Two reactions indicate clearly the presence of strain within the structure both by the products formed and the readiness with which the reactions take place. First, hydrolysis with 0.1 N sulphuric acid at room temperature gives 3, 6-anhydrogalactose, the properties of which indicate that it has an aldehydic structure rather than the usual hemiacetal i.e. that the pyranose ring has been opened (Fig. 1b) and second, treatment with excess 0.5% methanolic hydrogen chloride at room temperature again causes opening of the pyranose ring with the formation of the dimethyl acetal (Fig. 1c).

/The ...



a.

Methyl-3,6-anhydro-α-D-galactoside (angal)

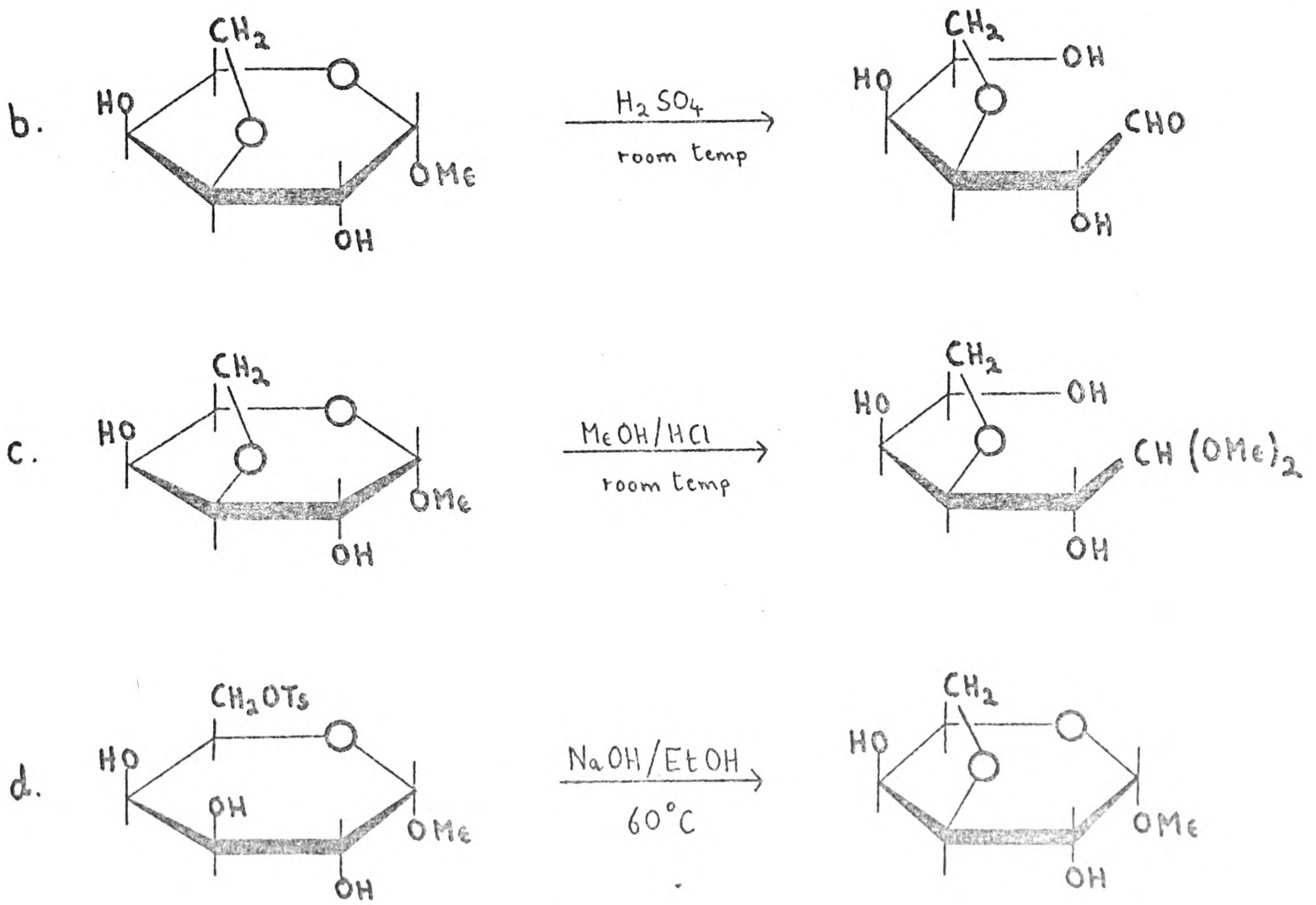


Fig. 1. The structure, reactions and preparation of methyl-3,6-anhydro-α-D-galactoside (angal).

The presence of strain within the dicyclic system influences the method of preparation of angal as the pyranose ring must be stabilised as the methyl glycoside before the 3, 6-anhydro ring is formed. In all the reactions studied by Haworth et al. the five membered ring remained intact.

1.2, PREPARATION OF ANGAL (1)

6-tosyl-α-methyl-galactoside (30g) in ethanol (1200 ml.) with NaOH (1N: 95 ml.) were heated at 60°C for one hour, neutralised with solid CO_2 and concentrated to dryness under reduced pressure. The solid residue was extracted with acetone until the extract gave a negative phenol/sulphuric acid test for carbohydrate material (2). The acetone solution was evaporated to dryness under reduced pressure and the product was recrystallised from alcohol. The crude yield was 10.6 g. The product was recrystallised three times in order to obtain crystals of high quality suitable for x-ray diffraction studies. The melting point of angal was 136-138°C.

1.3, CRYSTAL DATA

Angal crystallised as colourless needles of rectangular or six sided cross section. The crystals showed extinction parallel to the needle axis when viewed under the polarising microscope. Preliminary photographs were taken using a Unicam equi-inclination Weissenberg camera.

Needle axis ('c' axis) and 'a' axis oscillation and Weissenberg photographs showed that $I(\bar{h}kl) \equiv I(\bar{h}\bar{k}l) \equiv I(h\bar{k}l) \equiv I(hkl)$ and that the reflections h00, 0k0 and 00l were only present for

/even...

even values of h, k and l. The crystals were thus orthorhombic with space group $P2_1\ 2_1\ 2_1$. Cell dimensions were measured from charts and later from measurements of calibrated films and the density of the crystals was found by weighing a known volume of a mixture of carbon tetrachloride and benzene in which angal crystals remained suspended. The crystal data are summarised in Table 1.

Table 1: Angal Crystal Data

Space group	$P2_1\ 2_1\ 2_1$.
Cell dimensions (Å)	a= 9.459±0.010
	b= 12.049 ± 0.010
	c= 6.932 ±0.015
Cell volume $(Å)^3$	790.1
M.W.	176.3
Calculated density $(g.cm^{-3})$	1.482
Observed density $(g.cm^{-3})$	1.48
no.mols/unit cell	4
mass absorption coefficient μ (CuKα) (cm^{-1})	11.02

1.4, DETERMINATION OF CELL DIMENSIONS

Zero layer Weissenberg photographs, calibrated with silicon powder lines down each edge, were taken of angal crystals

/mounted ...

mounted about the 'c' axis and 'a' axis. The theta values of the silicon powder lines were calculated (a = 5.431 Å for silicon (3,4)) and graphs were drawn of the separation of the powder lines against theta. Measurements on the films were made using a travelling microscope. Theta values were measured for a number of high angle reflections of angal, the Cu $K\alpha_1$ and Cu $K\alpha_2$ reflections both being measured where possible. The 'best' values of the cell dimensions were calculated using two least squares procedures. In the first, the values of the cell dimensions were systematically varied and those which gave the minimum value of $\Sigma (\theta_{obs} - \theta_{calc})^2$ were found. In the second, a linear least squares procedure was used to solve for 'a', 'b' and 'c', the quantity minimised being $\Sigma \left(\frac{4.\sin^2\theta}{\lambda^2} - h^2.a^{*2} - k^2.b^{*2} - l^2.c^{*2}\right)^2$ The cell dimensions are given in Table 1 together with estimated errors which are approximately three times the standard deviations obtained from the linear least squares procedure.

1.5, THE INTENSITY DATA

The reflection data were collected using a Unicam equi-inclination Weissenberg camera. The x-ray tube was a Philips fine focus tube, with a copper target, operated at 35KV and 20mA. A nickel filter was used to remove unwanted Cu $K\beta$ radiation. Ilford Industrial G x-ray film was used throughout.

Equi-inclination Weissenberg photographs of layers zero to five were taken of an angal crystal, mounted about its 'c' axis, using multiple film packs. The crystal used was approximately cylindrical in shape with a radius of 0.013 cm.

/A ...

A standard intensity strip was prepared by taking a series of 5 degree zero layer Weissenberg exposures with the camera being moved along between each exposure. Exposures were taken for the following number of single traverses of the camera:-
1,2,3,4,5,6,7,8,9,10,12,14,16,18,20,22,24,28,32 and 36.

The most suitable row of spots on the film was chosen such that the reflection assigned an intensity of one was only just visible.

The intensities of the reflections of the six layers were estimated visually by comparison with the spots on the standard intensity strip. On the upper layer photographs the peak density was estimated rather than an integrated intensity so that a spot shape correction could easily be applied. Where the α_1 and α_2 reflections were resolved the α_1 reflection was measured and its value increased by 50%. Inter-film scaling factors within a layer were found by plotting graphs of the measured intensities on one film against those on the next film in the pack. The gradients of the best straight lines through the origin gave the scale factors. The number of independent reflections measured was 746.

Inter-layer scaling factors were found from some carefully timed exposures with half of each photograph being a zero layer Weissenberg photograph and the other half being an upper layer Weissenberg photograph. The raw intensities were corrected for Lorentz and polarisation factors (3) and the Phillips spot shape correction (5) applied. No correction was made for absorption effects.

SECTION II: PROJECTION WORK AND PATTERSON FUNCTIONS

2.1, TWO DIMENSIONAL PATTERSON FUNCTION (6)

A two dimensional Patterson function, P(x,y), was calculated using Beevers-Lipson strips (7) where,

$$P(x,y) = 4 \sum_{h=0}^{h_{max}} \sum_{k=0}^{k_{max}} F^2_{obs}(hk0) \cdot \cos(2\pi hx) \cdot \cos(2\pi ky)$$

A sharpened Patterson function was also calculated on the computer, the modification factor applied to the F^2obs values being,

$$M = \frac{1}{\hat{f}^2} \cdot (\sin\theta)^{m_1} \cdot \exp(m_2 \cdot \sin^2\theta/\lambda^2)$$

$\hat{f}$ = mean scattering factor at θ.

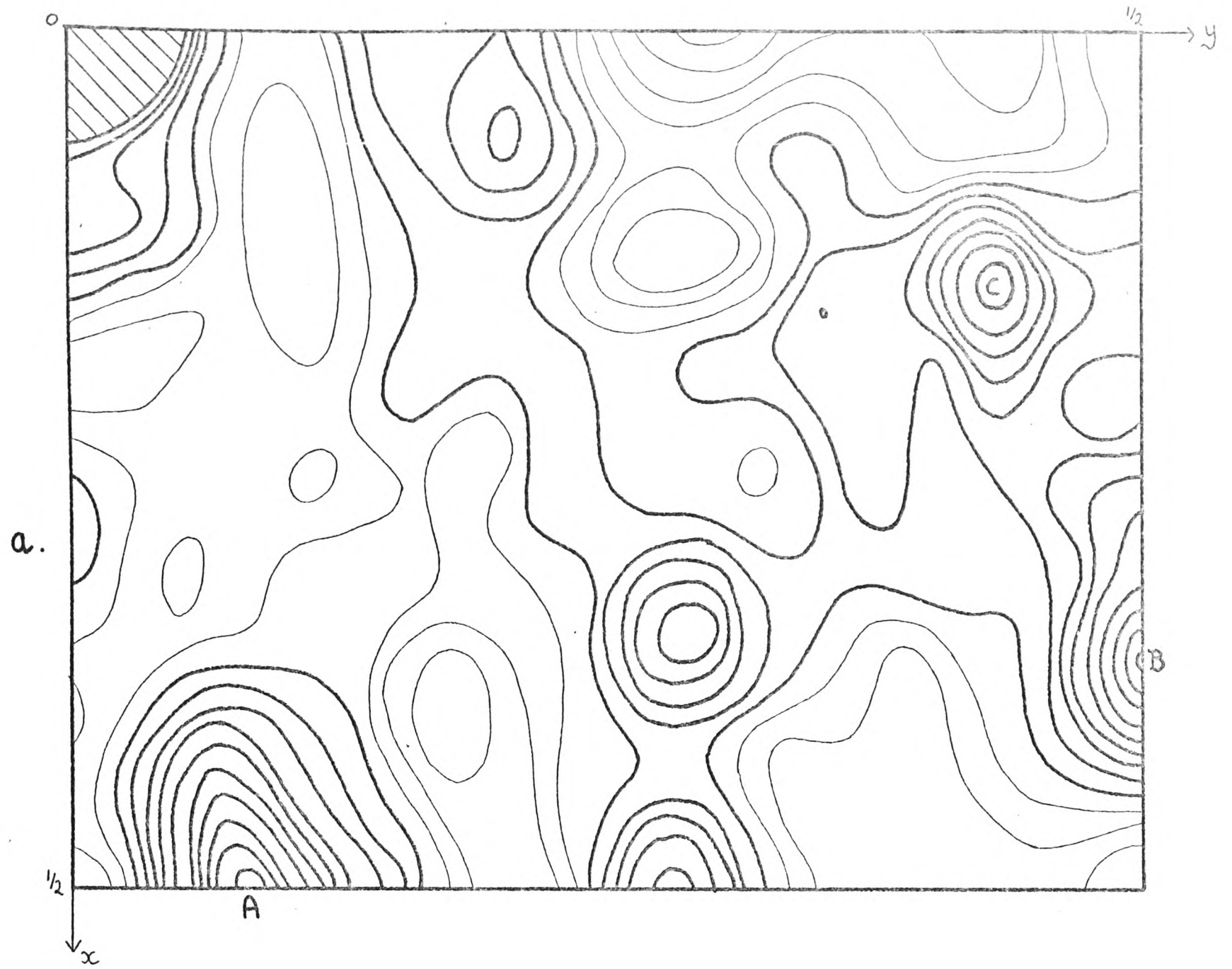
Only the temperature factor part of the sharpening function was used with $m_1 = 0$ and $m_2 = 8$ and with $\hat{f}^2 = 1$ for all values of θ. The Patterson function had symmetry Pmn and part of the sharpened function is shown in Fig. 2a.

A model was constructed to represent the multiple weight intra-molecular vectors of angal. Wires of lengths proportional to the vectors were inserted in a central plastic ball drilled at appropriate angles. Different coloured balls were placed on the free ends of the wires representing the different multiplicities. The most useful multiple weight vectors for determining the orientation of a pyranose ring are the 2.5 Å vectors across the ring (8). These are illustrated in Fig. 2b and give a planar set of six vectors as shown. There would be some distortion in the case of angal due to the effect of the 3, 6-anhydro ring on the conformation of the pyranose ring.

The vector model was projected onto the Patterson function

/using ...

a.



b.

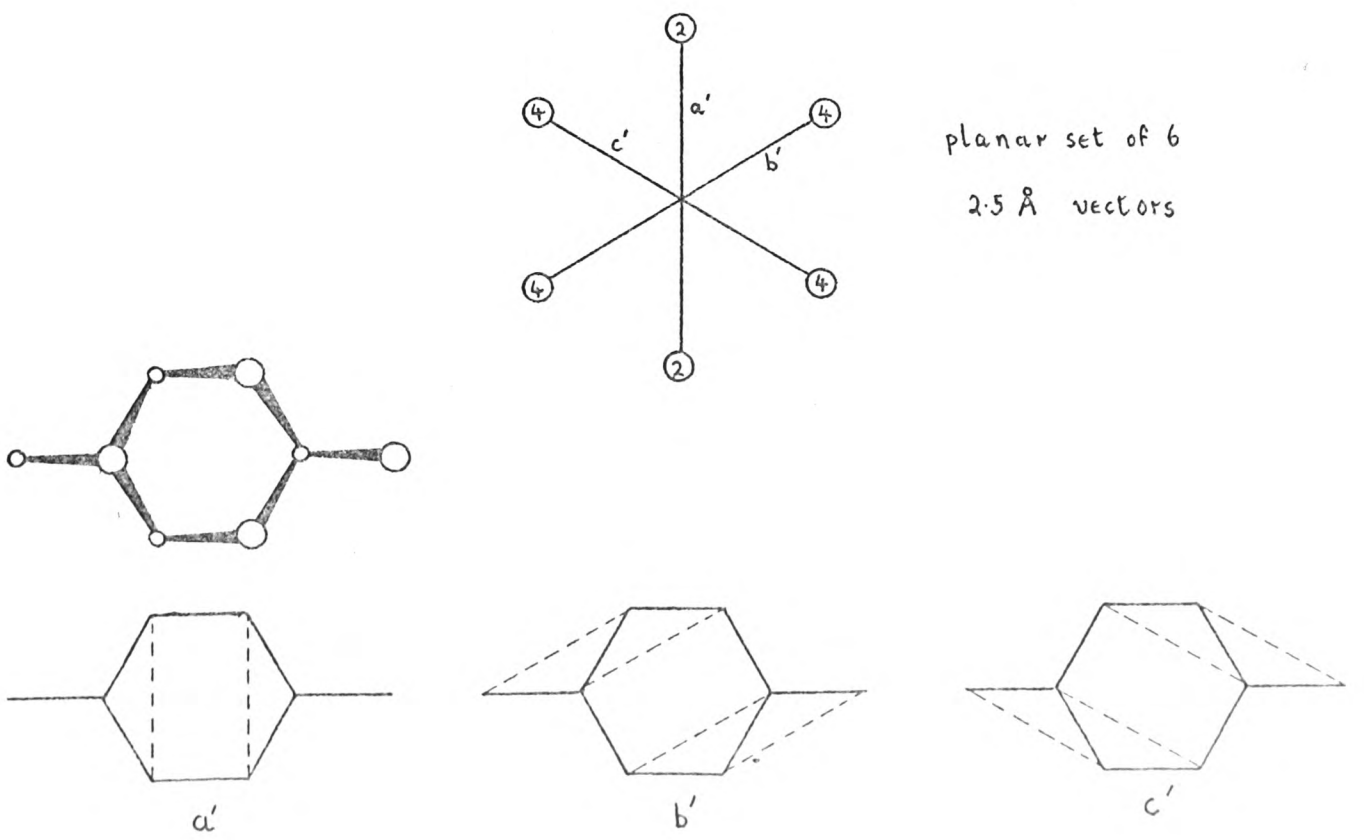


Fig. 2. (a) Part of the two dimensional Patterson function of angal (hk0). (b) The 2.5Å vectors of the pyranose ring.

using a parallel light beam. The centre of the model was placed over the origin of the Patterson function and the model was rotated to find the orientation which gave the best fit of the vectors to the peaks of the function. The best orientation appeared to be that with the pyranose ring tilted ap about 60 degrees to the xy plane and with the O(1)...O(4) direction approximately parallel to the y axis (Fig. 3a).

The two peaks A and B (Fig. 2a) on the Harker lines (10) of the Patterson function at $x = \frac{1}{2}$ and $y = \frac{1}{2}$ together with peak C were interpreted as being the multiple weight ring centre to ring centre vectors (8). The hko projection for space group $P2_1\ 2_1\ 2_1$ has symmetry pgg. If the origin of the projection is taken on a centre of symmetry (i.e. at x = 0.25, y = o with respect to the three dimensional origin) then the equivalent positions for the projections are

x; y

-x; -y

$\frac{1}{2}+x;\frac{1}{2}-y$

$\frac{1}{2}-x;\frac{1}{2}+y$

For a ring centred at XY this gives the three vectors A, B and C at $(\frac{1}{2};\ \frac{1}{2}-2Y)$, $(\frac{1}{2}-2X;\frac{1}{2})$ and (2X;2Y) respectively. This gave ring centre positions of angal at (0.567, 0.718), (0.433, 0.282), (0.067, 0.782) and (0.933, 0.218).

2.2, FURTHER WORK IN PROJECTION (hk0)

The information derived from the two dimensional Patterson function taken along with the study of some structure factor

/graphs ...

graphs, prepared for some of the strongest low order reflections, formed the basis of some trial structures of type a (Fig. 3). The best of these had an R factor ($= \Sigma \| |F|_{obs.} - |F|_{calc.} | / \Sigma |F|_{obs.}$) of 0.55.

The angal molecule was systematically rotated about the ring centre position derived from the Patterson function. $\Sigma (|F|_{obs.} - |F|_{calc.})^2$ was calculated for fifty reflections for each orientation of the molecule and the orientations which gave minimum values for this sum were found. The best of these were investigated further but none of them gave rise to a satisfactory solution of the structure. (See section 7.1)

Signs of 21 of the reflections with the largest normalised structure factors (E's) were derived from the triple product relationship (9) and a two dimensional Fourier synthesis was calculated with the signed E's as coefficients (E-map). Two more trial structures were proposed. These were of types b and c (Fig. 3). An attempt was made to refine these using a full matrix least squares procedure. In the first case (type b) the R factor was reduced to 0.27 but the atomic positions were no longer physically meaningful. The second structure did not refine (See section 7.4).

Fig.3 summarises the types of trial structure that were investigated in the projection work. The relative merits of the structures are discussed briefly below.

(i) Type a was consistent with the two dimensional Patterson function and gave fairly good low order structure factor agreement.

/(ii) Type b ...

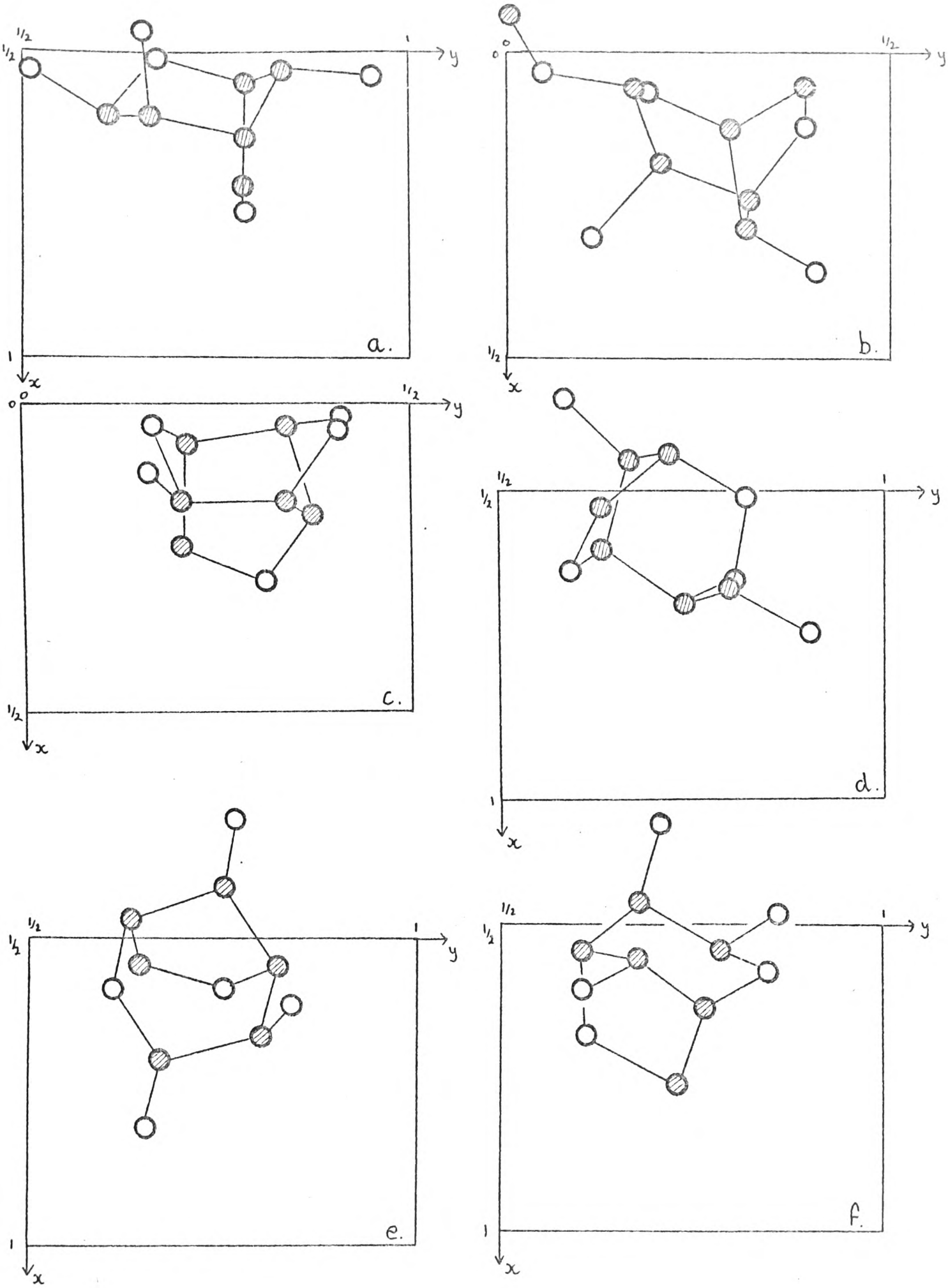


Fig. 3. Some trial structures from the hk0 projection work on the structure of angal.

(ii) Type b was consistent with the calculated set of 21 signs and gave good low order structure factor agreement. The orientation of the pyranose ring was not consistent with the Patterson function and the attempted refinement was unsatisfactory.

(iii) Type c was also consistent with the calculated set of 21 signs and the orientation of the ring was consistent with the Patterson function. The low order structure factor agreement was rather poor and for some of the high intensity reflections the calculated values of the structure factors were much greater than the observed values.

(iv) Types d and e were derived from the systematic rotation and type d gave good low order structure factor agreement. The ring orientations were not consistent with the Patterson function.

(v) Type f was also derived from the systematic rotation. The orientation of the pyranose ring was consistent with the Patterson function but low order structure factor agreement was rather poor. From packing considerations types c and f seemed less likely than the other structures as the length of the molecule lay approximately along the direction of the shortest axis.

As the projection work did not lead to the solution of the structure further discussion of it is left until later (see Section VII).

/2.3, The Three ...

2.3, THE THREE DIMENSIONAL PATTERSON FUNCTION (6)

A three dimensional sharpened Patterson function,

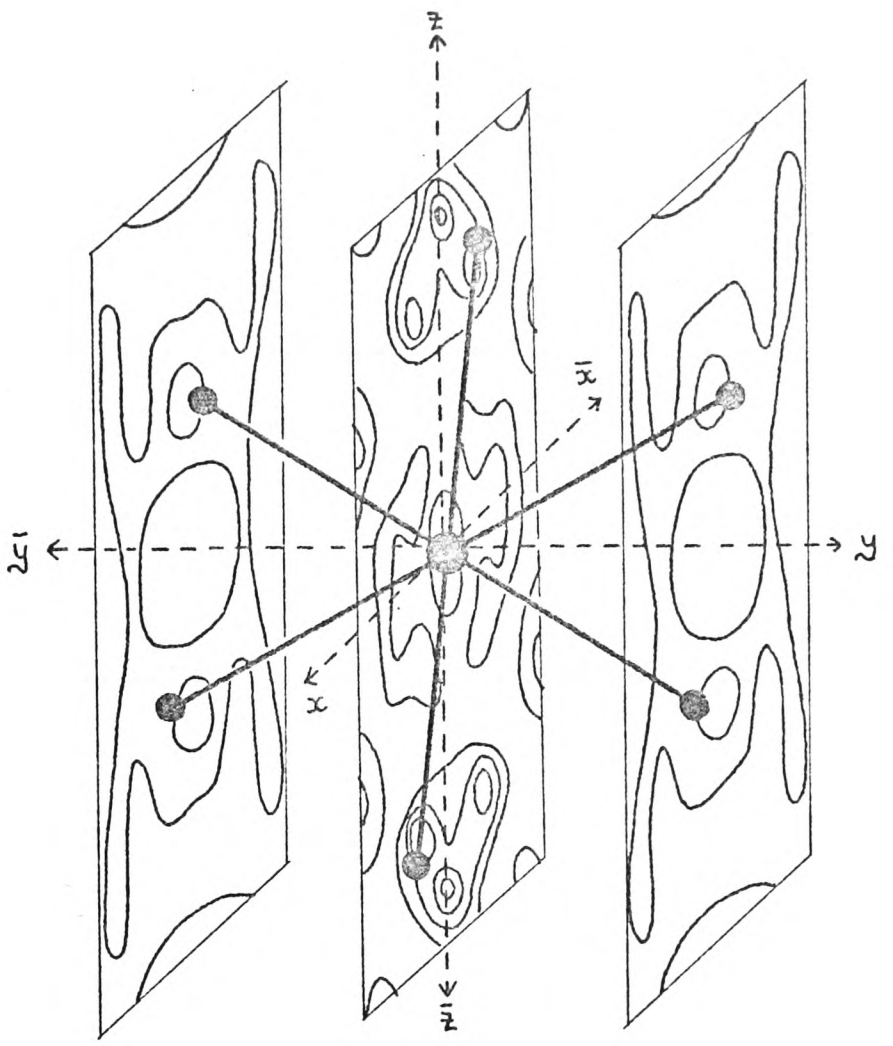
$$P(x,y,z) = \frac{1}{V} \cdot \sum_h \sum_k \sum_l F_s(hkl)^2_{obs.} \cdot \cos[2\pi(hx+ky+lz)]$$

was computed. The $F(hkl)^2_{obs.}$ values were sharpened as for the two dimensional Patterson function (Section 2.1). The symmetry of the Patterson function was Pmmm. Sections of the Patterson function along the y axis to a distance of about 3Å were plotted out on tracing paper and inspected for the 2.5 Å multiple weight vectors of the pyranose ring. Only one satisfactory solution was found for the orientation of the ring (Fig. 4a). In this the 'plane' of the ring lay approximately parallel to the y axis and inclined to the xy plane at an angle of about 60 degrees. The height of the Patterson function peak at 2.5 Å from the origin in the section at y = o, allowing for the fact that it was of double weight by virtue of lying on a mirror plane, suggested that the peak corresponded to one of the four weighted vectors rather than a double weighted vector. This suggested that O(1) and O(4) and the pyranose ring atoms were lying as shown in Fig. 4b.

As well as carrying out the visual inspection of the Patterson function, various rotation functions were computed to try and find the best orientation of the molecule. No significant new information was found. The rotation functions are described in Section 7.2.

The three Harker sections (10) of the Patterson function, at $x = \frac{1}{2}$, $y = \frac{1}{2}$ and $z = \frac{1}{2}$, were plotted out but no useful ihformation was found from them.

a.



b.

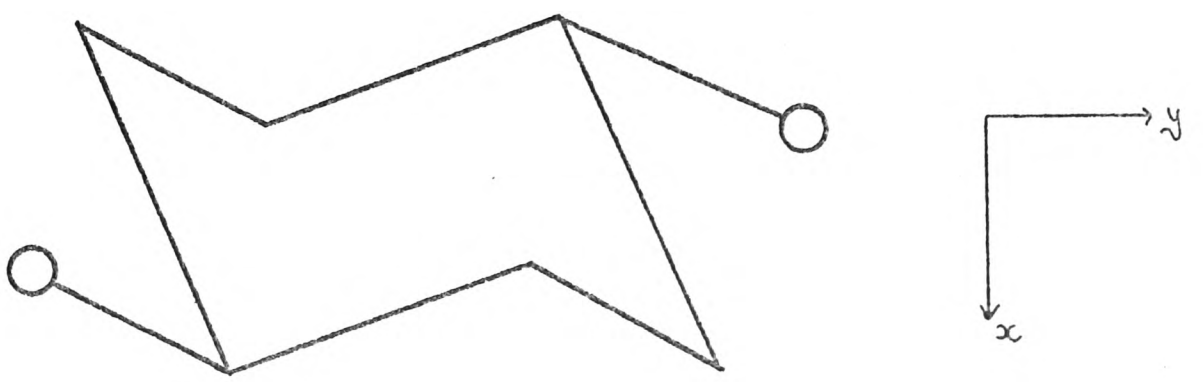


Fig. 4. (a) The fit of the 2.5Å vectors of the angal pyranose ring to the three dimensional Patterson function. (b) A possible ring orientation.

SECTION III: DIRECT METHODS, APPLICATION OF THE SYMBOLIC ADDITION PROCEDURE (11,12)

3.1, CALCULATION OF NORMALISED STRUCTURE FACTORS

Normalised structure factor magnitudes $|E|_h$ were calculated using the expression

$$|E|_h^2 = \frac{|F|_h^2}{\epsilon \cdot \sum_{j=1}^{N} f_j^2(h)}$$

where $|F|_h$ is the structure factor magnitude (on an absolute scale and temperature factor sharpened), f_j is the atomic scattering factor of the jth atom in a unit cell containing N atoms and ϵ is a number correcting for space group extinctions. In space group $P2_12_12_1$ $\epsilon = 2$ for reflections h00, 0k0 and 00l and $\epsilon = 1$ for all other reflections.

A Wilson plot (13) was calculated in order to find the scale factor required to put the observed structure factor magnitudes on an absolute scale and to obtain a value for the overall temperature factor. It has been recommended (14) that a K curve be used in the derivation of normalised structures as this does not involve making any assumptions about the nature of the crystal structure. In the present case, however, the Wilson plot was reasonably linear and was therefore considered to be satisfactory.

The average value of E^2 was 0.994. This was close enough to the theoretical value of 1.000 to make rescaling of the $|E|$ values unnecessary. The average value of $|E|$ was 0.885 compared with the theoretical value of 0.886 and of $|E^2 - 1|$ 0.729 compared with 0.736 (15).

/3.2, The Sigma-2....

3.2, THE SIGMA-2 LISTING

The basic formula for starting the phase determination was the sigma-2 formula:-

$$\phi_{\vec{h}} \doteq \langle \phi_{\vec{k}} + \phi_{\overrightarrow{h-k}} \rangle_{k_r}$$

where k_r implies that the average is only taken over the reflections with the larger values of $|E|$ and where $\phi_{\vec{h}}$, $\phi_{\vec{k}}$ and $\phi_{\overrightarrow{h-k}}$ are the phases of the reflections $\vec{h}$, $\vec{k}$ and $\overrightarrow{h-k}$.

A computer program was written to make a list (sigma-2 list) of all sets of three reflections $\vec{h}$, $\vec{k}$ and $\overrightarrow{h-k}$ where $|E_{\vec{h}}|$, $|E_{\vec{k}}|$ and $|E_{\overrightarrow{h-k}}|$ were all greater than 1.5. For angal there were 88 reflections with $|E|$ greater than 1.5.

Each of the 88 reflections (basic set, $\vec{h}$) was taken with each of the eight symmetry mates * of each reflection in the basic set (expanded set, $\vec{k}$) and $\overrightarrow{h-k}$ calculated. The expanded set was examined to see whether it contained the reflection $\overrightarrow{h-k}$ and, if it did, the set of three reflections $\vec{h}$, $\vec{k}$ and $\overrightarrow{h-k}$ was printed out. If $\overrightarrow{h-k}$ was not in the expanded set then the relationship was not printed out as $|E_{\overrightarrow{h-k}}|$ was less than 1.5 in this case.

To avoid having to search through the expanded set each time to find whether it contained the reflection $\overrightarrow{h-k}$, the reciprocal lattice was stored in a three dimensional array running from $-h_{max}$ to h_{max}, $-k_{max}$ to k_{max} and $-l_{max}$ to l_{max} where h_{max}, k_{max} and l_{max} were the maximum values of h, k and l for reflections in the basic set. Each reflection in the basic set was assigned a reference number i and this number was stored in the array location $h_i\ k_i\ l_i$ and also in $h_i\ k_i\ \bar{l}_i$, $h_i\ \bar{k}_i\ l_i$ etc. All other array locations were set to zero. Thus if a reflection $\overrightarrow{h-k}$ had indices h' k' l' it was only necessary to look up the array location h' k' l' to

* i.e. hkl, hk$\bar{l}$, h$\bar{k}$l, h$\bar{k}\bar{l}$, $\bar{h}$kl, $\bar{h}$k$\bar{l}$, $\bar{h}\bar{k}$l and $\bar{h}\bar{k}\bar{l}$

/to ...

to see if it was in the expanded set. If the number found at h' k' l' was a zero then the relationship was not valid. If there was any other number then a valid sigma-2 relationship had been found and the number found in the location was the reference number of the reflection in the basic set. This number was used to retrieve other information about the reflection such as its |E| value. The principle of the method for a two dimensional case is illustrated in Fig. 5.

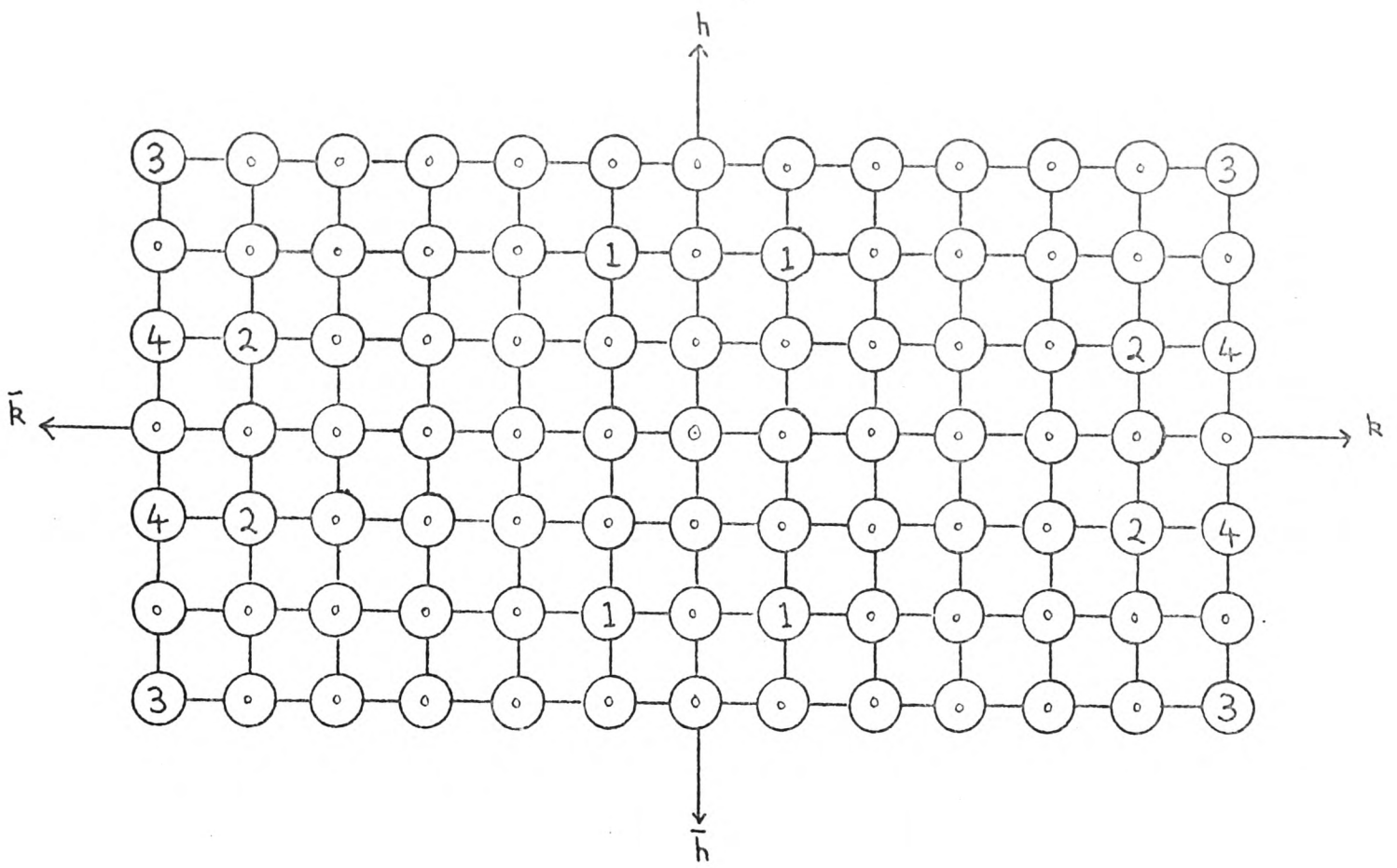
In using the sigma-2 formula it is necessary to know how the phase of a symmetry mate is related to that of the basic reflection. To avoid having to work this out for each sigma-2 relationship used, a routine was written so that the computer would print out the appropriate phase relationships alongside each sigma-2 relationship.

e.g. $\vec{h}$ $\vec{k}$ $\overrightarrow{h-k}$

2 1 0 1 $\bar{8}$ 3 1 9 $\bar{3}$ A PI-A -A

All $\vec{h}$ have positive indices and thus have the basic phase A. The (PI-A) indicates that the phase of the $1\bar{8}3$ reflection is equal to π minus the phase of the 183 reflection. Similarly the -A indicates that the phase of the $19\bar{3}$ reflection is minus that of the 193 reflection. The reference numbers of the three reflections, the product of the |E| values and the number of the sigma-2 interaction were also output. An extract from the sigma-2 list is given in Fig. 6. The number of sigma-2 relationships, into which each reflection entered, was counted and printed at the end of the list.

/3.3, Specification ...



Array: -hmax to +hmax, -kmax to +kmax

basic set,

i	$\vec{h}$	$\vec{k}$
1	2	1
2	1	5
3	3	6
4	1	6

examples,

	$\vec{h}$	$\vec{k}$	$\overrightarrow{h-k}$		
(1)	21	$\bar{1}\bar{5}$	36		valid (3 in location 3,6)
(2)	15	$\bar{1}6$	$21\bar{}$		valid (1 in location 2,$\bar{1}$)
(3)	21	$\bar{1}6$	$3\bar{5}$	not	valid (0 in location 3,$\bar{5}$)

Fig. 5. An illustration of the method of the igma-2 calculation.

3.3, SPECIFICATION OF ORIGIN (16)

In space group $P2_12_12_1$ the origin is specified by assigning values of phases to three linearly independent reflections. These reflections should have high values of $|E|$ and should enter into a large number of sigma-2 relationships. It is preferable to choose two dimensional reflections whose phases are either 0 or π or $\pm \pi/2$ depending on the parity of the indices. This ensures that the origin defined is a standard one for the space group.

In the present instance the number of reflections with large $|E|$ values was rather small and a compromise had to be reached in attempting to fulfil the conditions described above. The condition of linear independence was however imperative.

The three reflections, chosen to define the origin, together with the phases assigned to them, are listed below:-

h	k	l	ϕ	E	no. $\Sigma 2$ interactions
5	0	5	$\pi/2$	2.28	39
0	8	3	0	2.03	32
2	1	0	0	1.67	76

The $|E|$ value of the 210 reflection was rather lower than would normally be considered desirable* but it was possible to determine the phase of the 315 reflection using it. This reflection had $|E| = 2.43$ and entered into 73 sigma-2 interactions and was thus potentially a very useful phase.

3.4, PROBABILITY CONSIDERATIONS

The variance V in square radians of a phase determined from the known values of other phases using the sigma-2 relationship

* The measured intensity of the 210 reflection was in fact far too low and the correct value of E should have been about 2.52.

/may ....

may be expressed in terms of α where:-

$$\alpha = \left\{ \left[\sum_{kr} K_{(\vec{h},\vec{k})} \cdot \cos\left(\phi_{\vec{k}} + \phi_{\overrightarrow{h-k}}\right) \right]^2 + \left[\sum_{kr} K_{(\vec{h}\,\vec{k})} \cdot \sin\left(\phi_{\vec{k}} + \phi_{\overrightarrow{h-k}}\right) \right]^2 \right\}^{1/2}$$

where $K_{(\vec{h},\vec{k})} = 2 \cdot \sigma_3 \cdot \sigma_2^{-3/2} \cdot |E_{\vec{h}} \cdot E_{\vec{k}} \cdot E_{\overrightarrow{h-k}}|$

and $\sigma_n = \sum_{j=1}^{N} z_j^{\,n}$

N = no. atoms in unit cell.

z_j = atomic number of the jth atom.

A graph of V against α has been plotted by J. Karle and I.L. Karle (11).

Considering the case where there is only a single interaction to define a phase, the expression for α reduces to

$$\alpha = K_{(\vec{h},\vec{k})} = 2 \cdot \sigma_3 \cdot \sigma_2^{-3/2} \cdot |E_{\vec{h}} \cdot E_{\vec{k}} \cdot E_{\overrightarrow{h-k}}|$$

$\sigma_3 \cdot \sigma_2^{-3/2}$ is a constant for a given crystal and was 0.149 for angal.

A working rule for accepting a phase indication is that the variance should be less than 0.5. V = 0.5 corresponds to $\alpha \sim 2.5$ and $|E_{\vec{h}}, E_{\vec{k}}, E_{\overrightarrow{h-k}}| = 8.4$ for angal.

A phase indication was considered to be acceptable if

(i) The product of the three |E|'s was greater than 8.5. for a single indication.

or (ii) The sum of the product of the three |E|'s was greater than 10.0 for a number of similar indications for the phase.

3.5, USING THE SIGMA-2 LIST

In the sigma-2 list each sigma-2 relationship was printed out

/twice ...

twice for each $\vec{h}$. Thus the relationship between reflections 42, 45, 5 was printed out as well as the equivalent 42, 5, 45 (Fig. 6). Such duplications could easily be removed by imposing the condition that $i_{\vec{k}}$ should be greater than $i_{\overrightarrow{h-k}}$ but it was convenient to leave them in for carrying out a systematic exploration of the sigma-2 list for phase indications. Once a phase had been assigned a value, its reference number was underlined wherever it occurred in the $i_{\vec{k}}$ and $i_{\overrightarrow{h-k}}$ columns. As the $i_{\vec{k}}$ was arranged in increasing order of magnitude of i for each $\vec{h}$, it was easy to find when a given reflection occurred. Wherever adjacent reflections were underlined in the $i_{\vec{k}}$ and $i_{\overrightarrow{h-k}}$ columns there was an indication of the phase in the $i_{\vec{h}}$ column. For example $\phi(4, 0, 13) = \phi(1, 3, 14) - \pi - \phi(5, 3, 1)$ from sigma-2 relationship number 1310 with the product of the three |E| values = 11.23. (Fig. 6). ($\phi(h,k,l)$ = phase of reflection (hkl)).

Each phase indication was noted as it was found together with the product of the three |E|'s for that indication.

3.6, LISTINGS OF PHASE INDICATIONS

During the phase determination, three listings were made apart from the sigma-2 list. These were

(i) The primary list, consisting of phases which had been accepted. When a phase was added to this list all occurrences of the reflection were underlined in the sigma-2 list. All new phase indications were noted before another phase was added to the list. Once a phase had been accepted, no further indications of its value were looked for.

/(ii) The ...

4	0	13	2	4	2	2	-4	11	6.004	42	5	45	1302	A	A	PI-A
4	0	13	2	-4	2	2	4	11	6.004	42	5	45	1303	A	-A	A
4	0	13	2	5	13	2	-5	0	7.303	42	12	56	1304	A	A	PI-A
4	0	13	2	-5	13	2	5	0	7.303	42	12	56	1305	A	-A	A
4	0	13	5	4	-1	-1	-4	14	5.616	42	25	74	1306	A	-A	A-PI
4	0	13	5	-4	-1	-1	4	14	5.616	42	25	74	1307	A	A-PI	PI-A
4	0	13	-1	0	15	5	0	-2	5.475	42	28	73	1308	A	PI-A	PI-A
4	0	13	-1	3	14	5	-3	-1	11.232	42	31	79	1309	A	-A	A
4	0	13	-1	-3	14	5	3	-1	11.232	42	31	79	1310	A	A-PI	-A
4	0	13	2	4	11	2	-4	2	6.004	42	45	5	1311	A	A	-A
4	0	13	2	-4	11	2	4	2	6.004	42	45	5	1312	A	PI-A	A
4	0	13	2	5	0	2	-5	13	7.303	42	56	12	1313	A	A	-A
4	0	13	2	-5	0	2	5	13	7.303	42	56	12	1314	A	PI-A	A
4	0	13	5	0	-2	-1	0	15	5.475	42	73	28	1315	A	PI-A	PI-A
4	0	13	-1	4	14	5	-4	-1	5.616	42	74	25	1316	A	PI-A	A-PI
4	0	13	-1	-4	14	5	4	-1	5.616	42	74	25	1317	A	A-PI	-A
4	0	13	5	3	-1	-1	-3	14	11.232	42	79	31	1318	A	-A	A-PI
4	0	13	5	-3	-1	-1	3	14	11.232	42	79	31	1319	A	A	-A

2	4	7	4	0	10	-2	4	-3	5.802	43	1	48	1320	A	A	A-PI
2	4	7	0	8	-2	2	-4	9	4.756	43	4	52	1321	A	-A	PI-A
2	4	7	-2	4	-2	4	0	9	5.265	43	5	46	1322	A	A	A
2	4	7	-2	-4	-2	4	8	9	5.863	43	5	44	1323	A	-A	A
2	4	7	0	-1	-5	2	5	12	7.641	43	6	56	1324	A	A-PI	A
2	4	7	2	-3	9	0	7	-2	7.368	43	8	57	1325	A	-A	-A
2	4	7	2	3	13	0	1	-6	6.288	43	9	59	1326	A	A	-A
2	4	7	2	5	13	0	-1	-6	5.550	43	12	59	1327	A	A	A-PI
2	4	7	-3	2	13	5	2	-6	5.953	43	17	76	1328	A	PI-A	PI-A
2	4	7	-3	-2	13	5	6	-6	7.863	43	17	67	1329	A	A	PI-A
2	4	7	-1	2	15	3	2	-8	6.023	43	18	71	1330	A	PI-A	PI-A
2	4	7	5	4	-1	-3	0	8	6.386	43	25	68	1331	A	-A	PI-A

Fig. 6. An extract from the angal sigma-2 list.

(ii) The secondary list, consisting of phase indications which were acceptable by the criteria described in Section 3.4. By making this listing, it was possible to choose, for the next addition to the primary list, the phase that had the strongest indications at that stage. As the phase indications which met the required criteria were not immediately accepted, it was often possible to get further indications of the phase, either giving increased certainty of its correctness or, if the new phase indication was contrary to the previous ones, decreasing the reliability of the previous indications. In the latter case the phase was sometimes removed from the secondary list if the contrary indications were sufficiently strong.

(iii) Tertiary list, consisting of each phase indication found together with the product of the three $|E|$'s for that indication.

Though the sigma-2 formula indicates that the average value of the phase indications should be taken, in practice, with a few exceptions towards the end of the phase determination, the majority indication was taken. This meant that the phases were all assigned values of 0, π or $\pm\pi/2$.

Three examples of indications which led to the acceptance of a phase are given below. (Eprod = $|(E_{\vec{h}} . E_{\vec{k}} . E_{\overrightarrow{h-k}})|$.)

(i) Reflection number 8 (392); phase = $\pi/2$

1 indication of $\pi/2$, Eprod = 10.8.

(ii) Reflection number 85 (133), phase = π - c*

* c was the symbol assigned to the 150 reflection (see section 3.9)

/3 indications ...

3 indications of π - c, Eprod = 7.2

π - c, Eprod = 7.5

π - c, Eprod = 4.9

Sum of Eprod's = 19.6

(iii) Reflection number 71 (283); phase = 0

11 indications of 0, Eprod = 5.8; 0, Eprod = 6.4

0, Eprod = 5.9; 0, Eprod = 4.9

π, Eprod = 5.9; 0, Eprod = 6.1

0, Eprod = 6.5; π, Eprod = 5.5

π, Eprod = 5.1; 0, Eprod = 6.5

0, Eprod = 6.5;

Sum of Eprod's for phase = 0 = 48.6 (8 indications)

Sum of Eprod's for phase = π = 16.5 (3 indications)

3.7, DERIVATION OF SOME PHASES

From the origin specifying phases, some further phases were found using the sigma-2 formula. Two examples are given below.

(i)

$\vec{h}$	$\vec{k}$	$\overrightarrow{h-k}$			
293	210	083	A	A	A

Phase (293) = phase (210) + phase (083) = 0 + 0 = 0

(ii)

$\vec{h}$	$\vec{k}$	$\overrightarrow{h-k}$			
315	505	$\bar{2}10$	A	A	PI - A

Phase (315) = phase (505) + phase ($\bar{2}10$)

= phase (505) + π - phase (210)

= $\pi/2 + \pi - 0$

= $3\pi/2$

During the phase determination all phases were increased or

/decreased ....

decreased by multiples of 2π such that /phase/ $\leq \pi$

Thus phase (315) = $3\pi/2 - 2\pi = -\pi/2$ (Eprod = 9.3)

Nine phases were found in terms of the three origin specifying phases.

3.8, SPECIFICATION OF ENANTIOMORPH

The enantiomorph is specified by fixing the value of a phase which is a structure in-variant. If the structure in-variant has a phase ϕ for one enantiomorph then it will have a value of $\pi+\phi$ for the other enantiomorph. By choosing one of these two phases the enantiomorph is specified, though it is not possible to tell at this stage which one it is. The value of the structure in-variant used must not be 0 or π and preferably should not be near to one of these values. In space group $P2_1\ 2_1\ 2_1$ the structure in-variants are those reflections with all three indices even. It is generally satisfactory to assign a value of $\pi/2$ to one of these structure in-variants provided that all three indices are non zero. (A two dimensional phase would have a value of 0 or π and could therefore not be used). It is however sometimes possible to define the value of a structure in variant by assigning a value to a two dimensional phase which is not itself a structure in-variant (17).

In specifying the enantiomorph it is desirable to fix the phase of a structure in variant which has a large value of |E| and which enters into a large number of sigma-2 interactions. The only reasonable choice for angal was the 2, 10, 2 reflection

/(E = 1.70 ....

($|E|$ = 1.70, number of sigma-2 interactions = 37). This reflection was assigned a symbol s. One further phase was determined in terms of this symbol.

3.9, ASSIGNMENT OF FURTHER SYMBOLS AND DETERMINATION OF MORE PHASES

At this stage a further symbol had to be assigned to a phase in order to proceed with the phase determination. The 150 reflection ($|E|$ = 2.14, number of sigma-2 interactions = 55) was assigned a symbol c where c could be either $\pm\pi/2$. Forty phases were determined in terms of the three origin specifying phases, s and c. The symbol d (= 0 or π) was assigned to the 502 reflection and the list of phases was expanded to 56. There were fairly strong indications that d was equal to π and it was given this value. There were also indications that s was $\pm\pi/2$ and it was set equal to $\pi/2$, thus specifying the enantiomorph.

Thus 60 phases were determined in terms of a single unknown symbol c. Setting c equal to $\pi/2$ and $-\pi/2$ in turn gave two possible sets of 60 phases. For the set with c = $\pi/2$, two further phases were assigned values.

3.10, TANGENT FORMULA REFINEMENT OF PHASES

Sets of approximate phases, derived for example by the symbolic addition procedure, may be refined using the tangent formula (17).

$$\tan \phi_{\vec{h}} \simeq \frac{\sum_{\vec{k}} |E_{\vec{k}} \cdot E_{\vec{h}-\vec{k}}| \cdot \sin(\phi_{\vec{k}} + \phi_{\vec{h}-\vec{k}})}{\sum_{\vec{k}} |E_{\vec{k}} \cdot E_{\vec{h}-\vec{k}}| \cdot \cos(\phi_{\vec{k}} + \phi_{\vec{h}-\vec{k}})}$$

/In contrast ...

In contrast with the sigma-2 formula the tangent formula is valid for all values of |E|

A computer program was written to carry out the tangent formula refinement. The structure of the program was similar to that of the sigma-2 listing program.

Tangent formula refinement is repeated until convergence is achieved. Usually two or three cycles are sufficient. As an indication of how a refinement is proceeding, unscaled values for $|E_{\vec{h}}| \cdot \cos\phi_{\vec{h}}$ and $|E_{\vec{h}}| \sin\phi_{\vec{h}}$ can be calculated (11):-

$$|E_{\vec{h}}| \cdot \cos\phi_{\vec{h}} \simeq \sigma_2^{3/2} \cdot \sigma_3^{-1} \cdot \langle |E_{\vec{k}} \cdot E_{\vec{h-k}}| \cdot \cos(\phi_{\vec{k}} + \phi_{\vec{h-k}}) \rangle_{k_r}$$

$$|E_{\vec{h}}| \cdot \sin\phi_{\vec{h}} \simeq \sigma_2^{3/2} \cdot \sigma_3^{-1} \cdot \langle |E_{\vec{k}} \cdot E_{\vec{h-k}}| \cdot \sin(\phi_{\vec{k}} + \phi_{\vec{h-k}}) \rangle_{k_r}$$

from which $$|E_{\vec{h}}|^2_{calc} = |E_{\vec{h}}|^2 \cdot \cos^2\phi_{\vec{h}} + |E_{\vec{h}}|^2 \cdot \sin^2\phi_{\vec{h}}$$

The $|E_h|_{calc}$ values are scaled to the $|E_h|_{obs}$ values such that $$\Sigma |E_{\vec{h}}|^2_{calc.} = \Sigma |E_{\vec{h}}|^2_{obs.}$$

and an R' factor defined by,

$$R' = \frac{\sum_{k_r} ||E_{\vec{h}}|_{obs} - |E_{\vec{h}}|_{calc.}|}{\sum_{k_r} |E_{\vec{h}}|_{obs}}$$

is calculated.

During the tangent formula refinement phases were omitted from further cycles of the refinement if the $|E_h|_{calc}$ values were very low (less than 0.3 to 0.5) or if only a few terms contributed to a phase. Phases which varied a large amount from cycle to cycle in a random manner were also taken out.

The two sets of 60 phases, derived from the symbolic addition procedure, were refined using the tangent formula. The

/version ...

The ~~version of the tangent~~ formula program used for this refinement did not check the numbers of interactions or the $|E|_{calc}$ values for the phases as described above. The first set with $c = \pi/2$ gave $R' = 0.168$ after four cycles of refinement and the second set with $c = -\pi/2$ gave $R' = 0.181$ after four cycles of refinement. In both cases convergence was virtually achieved after two cycles.

3.11, E-MAPS BASED ON 60 PHASES

Three dimensional E-maps (Fourier synthesis with $|E|$ or in this case $100\times|E|$ instead of $|F|$ as coefficients) were calculated for each set of 60 phases. Models of the E-maps were made using plasticene balls on wire rods to represent the positions of the peaks in the E-maps. For the first set of sixty phases the fifteen largest peaks were represented in the model. Two projections of the map are shown in Fig. 7.

The six peaks 5, 10, 12, 11, 6 and 3 formed a six membered ring in the boat conformation with all bond lengths close to 1.5 Å. It was considered to be unlikely that the six membered ring would be in a boat conformation and it was assumed that one of the peaks 3 or 12 was spurious. Inspection of the other peaks suggested a possible structure for the molecule with peaks corresponding of eight of the twelve non-hydrogen atoms of angal. The orientation of the ring was consistent with the three dimensional Patterson function and the structure was consistent with some of the projection work. It had been considered unlikely because of bad agreement between observed and calculated structure factors for the low order, high intensity reflections. A set of structure factors was calculated for an eleven atom trial structure corresponding to the left handed enantiomorph (all non-hydrogen atoms except the carbon atom of the methyl group). The R factor ($= \Sigma ||F|_{obs} - |F|_{calc}| / \Sigma |F|_{obs}$.)

/was ....

* The observed values of $|F|$ were scaled such that $\Sigma |F|_{obs.} = \Sigma |F|_{calc.}$

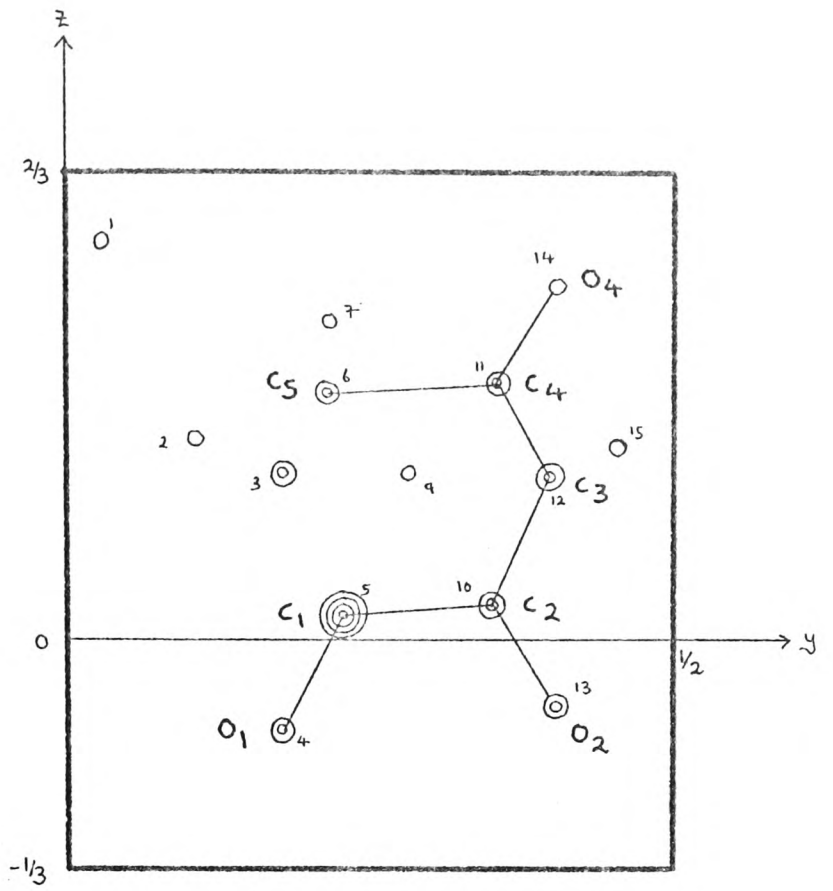


Fig. 7. Two projections (0kl, x=0 to x=$\frac{1}{2}$ and h'-0) of E-map based on first set of 60 phases.

was 0.49 and the structure factor agreement at low sin(θ) was poor. On looking through the list of structure factors, it was apparent that the calculated values of the structure factors were very much higher than the observed values for the high intensity reflections, suggesting that there was a systematic error in the measurement of these intensities. At this stage the trial structure did not seem sufficiently certain for a refinement to be attempted.

An E-map was calculated using the second set of sixty phases. It appeared to be less promising than the first E-map though it was possible to pick out a similar fragment of the angal molecule (Fig. 8). The ring was in a different symmetry position in the unit cell and the bond lengths and angles were less satisfactory.

3.12, EXPANSION OF THE PHASE LIST USING THE TANGENT FORMULA

Starting with the first set of 60 phases, the tangent formula was used to calculate additional phases and to refine them. 132 reflections with |E| values from 1.49 to 1.18 were included in the calculation. A phase was not accepted if there were fewer than three contributors or if the scaled $|E|_{calc}$ value was less than 0.4. The basic set of 60 phases was kept constant and only the newly determined phases were refined. Three cycles of refinement were carried out giving a total of 136 acceptable phases and an R' factor of 0.258. The addition and refinement of the phases was also carried out by Dr. W.E. Scott using his own tangent formula program. In this case all the phases were refined together. |E| values down to 1.20 were

/used ...

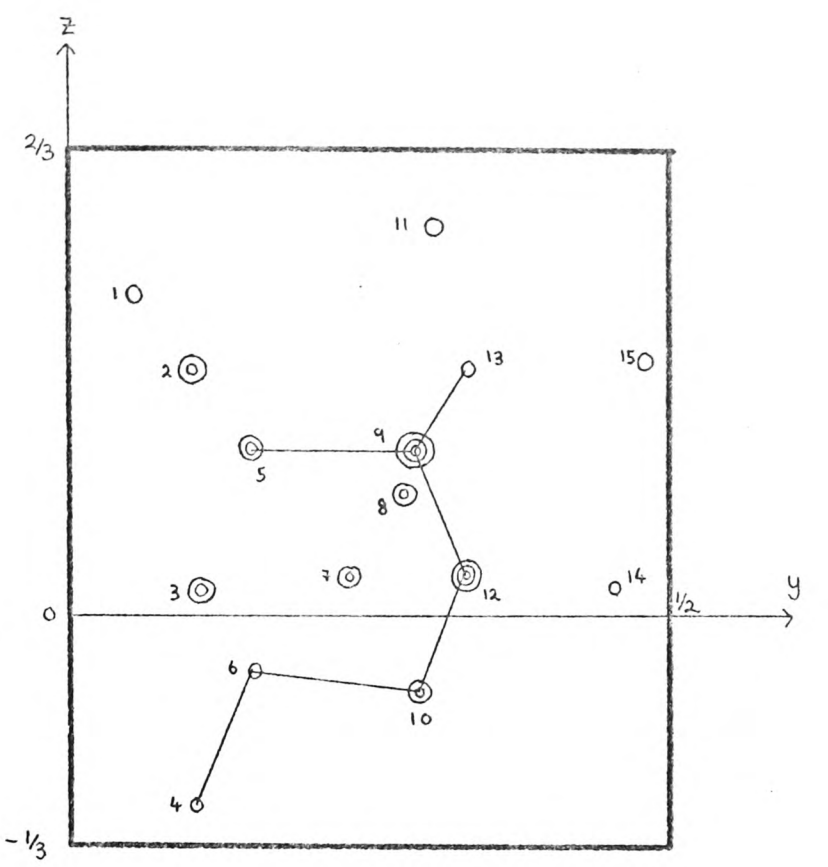


Fig. 8. Two projections (0kl, x=0 to x=$\frac{1}{2}$ and hk0) of E-map based on second set of 60 phases.

used and no phases were eliminated till the refinement was complete. At this stage phases, with low $|E|_{calc}$ values or which varied by a large amount during the refinement, were eliminated. The R' value was 0.297 and the total number of phases obtained was 120.

Three dimensional E-maps were calculated for the sets of 136 and 120 phases. No major differences were found between them and projections showing the major peaks are shown in Fig. 9.

The main difference between the 136/120 phase E-maps and the 60 phase E-map was the appearance of the peak labelled 1 which enabled a six membered ring in the chair conformation to be picked out. Peak 2, which turned out to be spurious, was still larger than peak 1. The heights of the significant peaks were greater than in the 60 phase E-map.

3.13, SEVEN ATOM PARTIAL TRIAL STRUCTURE

Structure factors and a Fourier synthesis were calculated from a seven atom partial trial structure based on the 136/120 phase E-maps. The six ring atoms were all treated as carbon atoms and an oxygen atom was placed at peak 3 of the E-map. No atom was placed at peak 10 as its position was probably not quite correct for O (2). It was hoped that this atom would show up in the Fourier synthesis, giving evidence of the correctness of the partial structure. The Fourier synthesis was computed with those terms for which $|F|_{calc}$ was greater than $0.4 \times |F|_{obs}$. The O(2) atom did not show up but there were two small peaks which were later found to correspond

/to ....

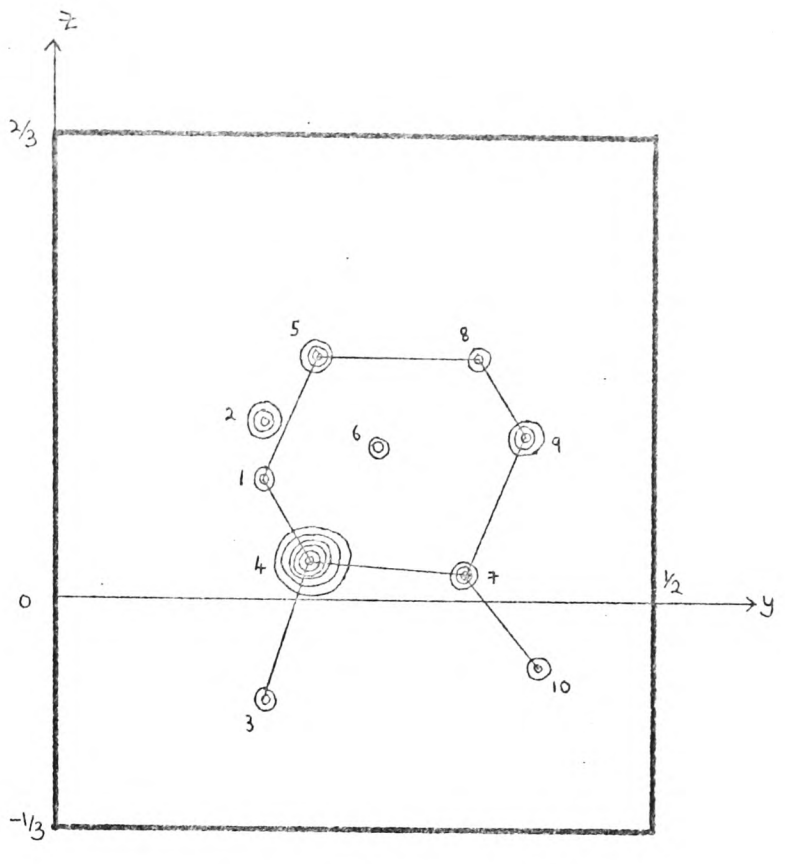


Fig. 9. Two projections (0kl, x=0 to x=$\frac{1}{2}$ and hk0) of E-map based on sets of 120 and 136 phases.

to O(3) and O(4). Both of these were slightly off the correct positions. The R factor for the seven atom partial trial structure was 0.50.

SECTION IV: COMPLETION AND REFINEMENT OF THE STRUCTURE

4.1, TANGENT FORMULA REFINEMENT OF PHASES DERIVED FROM SEVEN ATOM PARTIAL TRIAL STRUCTURE (18)

The criterion suggested for accepting an approximate phase derived from a partial trial structure is that $|F|_{calc} \geq p \times |F|_{obs}$ where p is the fraction of the total x-ray scattering power of the molecule contained in the partial structure. The $|E|_{obs}$ value of the reflection should be greater than 1.5. The seven atom partial trial structure described above represented just over half the scattering matter of the molecule but p was set equal to 0.4 to enable a few extra phases to be found. A set of 52 phases with $|E| \geq 1.5$ was derived. These phases were refined using three cycles of the Tangent formula ($R' = 0.139$). The list of phases was expanded using the tangent formula with 169 further reflections ($|E| \geq 1.17$) included in the calculation. After three cycles of refinement, eliminating after each cycle reflections with scaled $|E|_{calc}$ values less than 0.4 or which had less than three contributors, a set of 145 phases was obtained ($R' = 0.240$). In this case all the phases, including the basic set of 52, were refined together.

4.2, E-MAP FROM PHASES DERIVED FROM SEVEN ATOM PARTIAL STRUCTURE

An E-map was computed using the 145 phases derived from the seven atom partial trial structure. Two projections of the map are shown in Fig. 10. The three oxygen atoms O(2), O(3), and O(4) showed up as strong peaks in the E-map (nos. 11, 1 and 10) and there was a smaller peak corresponding to C(7)

/(no. 13)....

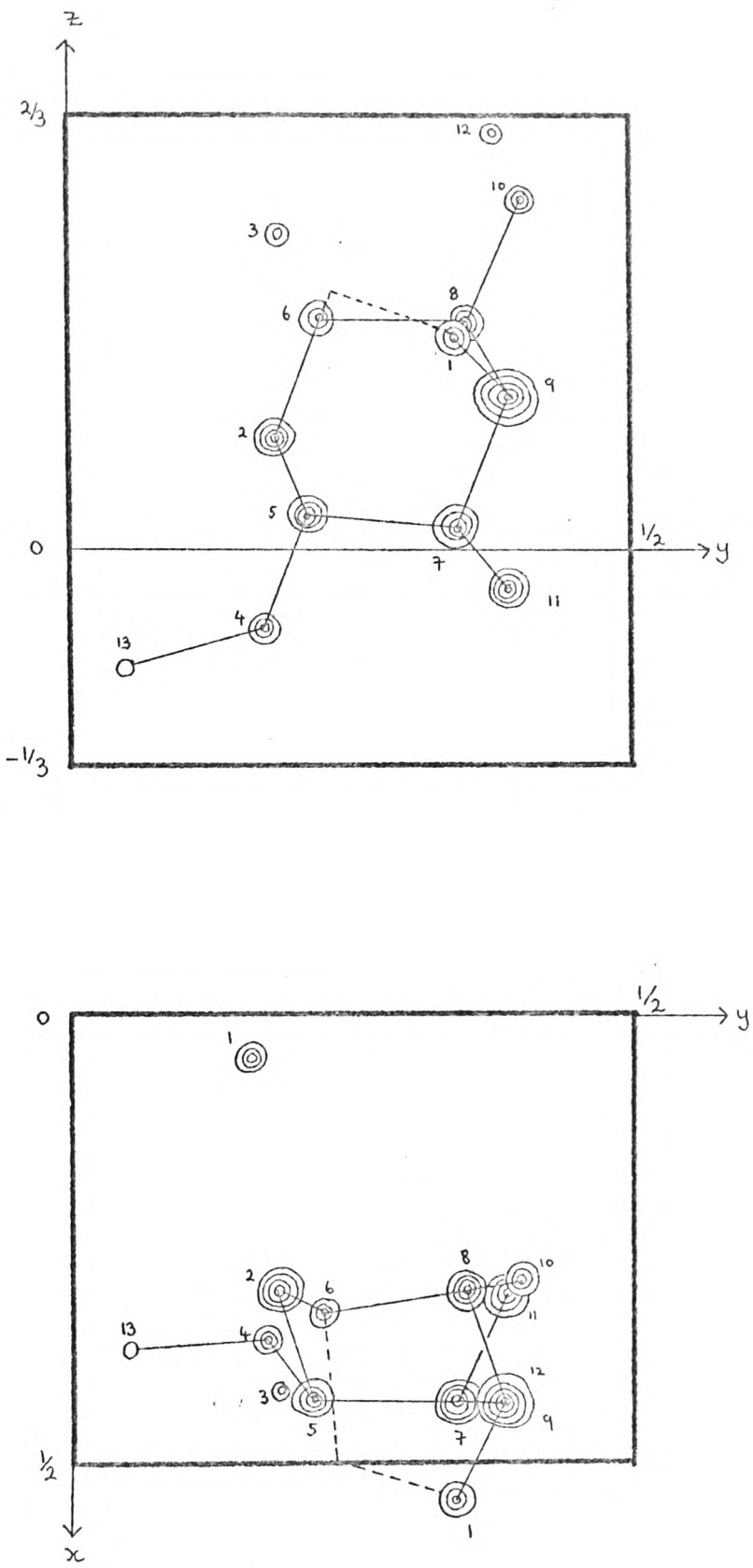


Fig. 10. Two projections (0kl, x=0 to x=½ and hk0) of E-map based on phases derived from seven atom partial trial structure.

(no. 13). There were also two spurious peaks (nos. 3 and 12). The positions of O(3) and O(4) were more satisfactory than those found in the Fourier synthesis. This fact taken along with the appearance of peaks corresponding to O(2) and C(7) illustrates the power of the tangent formula as a method for completing partial structures. As a further test of the method the process was repeated using a four atom partial trial structure. This is described in section 7.5.

4.3, COMPLETION OF THE STRUCTURE

Structure factors and a Fourier synthesis were calculated for a ten atom partial structure (all non-hydrogen atoms except C(6) and C(7)). The methyl group carbon atom was omitted as the peak in the E-map was fairly small and as its position was not closely defined by the rest of the structure. The R factor was 0.38. Terms with $|F|_{calc} \geq 0.4 \times |F|_{obs}$ were included in the calculation of the Fourier synthesis. The positions of C(6) and C(7) were easily identified. Structure factors were calculated for the complete trial structure (R = 0.26).

The appearance of all the atoms in chemically sensible positions and the rapid decrease in the R factor as the last few atoms were added left no doubt as to the correctness of the structure. The agreement between observed and calculated structure factors at low $\sin\theta$ was still rather poor and there was an obvious systematic error in the measurement of some of the large intensities where the $|F|_{calc}$ values were very much

/larger ...

larger than the $|F|_{obs}$ values. The effect was too large to be due to secondary extinction.

For example,

$|F(210)|_{calc}$ = 1194 , $|F(210)|_{obs}$ = 791

$|F(020)|_{calc}$ = 918 , $|F(020)|_{obs}$ = 604

In both these cases the observed value of the intensity was only about 0.44 of the calculated value.

4.4. POSSIBLE SOURCES OF ERROR IN MEASUREMENT OF HIGH INTENSITY REFLECTIONS

Possible sources of error in the measurement of the high intensity reflections were:-

(i) A systematic error in the visual estimation of the intensities.

(ii) Errors in the standard intensity strip. These could only have been small as a microdensitometer trace of the intensity strip showed that it was reasonably satisfactory.

(iii) Errors due to unsatisfactory film or unsatisfactory development of the films.

(iv) Inaccurate scale factors between the last few films of the film pack where there were only a few reflections, measured on each film, from which to obtain scale factors.

Of the sources of error listed above, number (iii), was probably the most important. The measurements and scaling factors for the last few films were checked for the zero layer

/film ...

film pack (the worst errors were in the zero layer reflections) and it appeared that the errors were on the films and not in the measurement of the films.

4.5, LEAST SQUARES REFINEMENT

The angal structure was refined using a full matrix least squares procedure (9). The quantity minimised was,

$$R_1 = \sum_i w_i \left[|F|_{obs} - k.|F|_{calc} \right]^2_i$$

where w_i was the weight of the ith reflection = $1/\sigma_i^2$, σ_i being its standard deviation.

The factor k was introduced as the $|F|_{obs}$ values were not on an absolute scale and the best value of k was found from the least squares. The refinement was carried out on a KDF9 computer using an Atlas Autocode program written by G.S. Pawley based on Busing, Martin and Levy's ORFLS (22).

4.6, WEIGHTING SCHEME

The weighting scheme used in the computer program was

$$W = 1/\sigma^2_{(F)}$$

where

$$\sigma(F) = (|F|/W3).W4 \quad \text{if} \quad |F| > W3$$

$$\sigma(F) = (W3/|F|).W4 \quad \text{if} \quad |F| < W3$$

$$\sigma(F) = W3.W4 \quad \text{if} \quad |F| < 1$$

|F| = observed value of structure factor magnitude.

W3 and W4 are constants whose values are input.

The highest weighted reflections are those with |F| = W3.

/For ...

For a set of intensities measured from films $(W3)^2$ is taken as having approximately the value of an intensity which was measured, on the top film of a film pack, in the most accurate region of the standard intensity strip. In the first stages of the refinement W3 was taken as 8 corresponding to a measured intensity of about 10 on the intensity strip. In the later stages, W3 was taken as 12 corresponding to a measured intensity of about 22. (both values of W3 have been placed on an absolute scale). W4 was used to put the standard deviations on an approximately absolute scale. For the highest weighted reflections with $|F| = W3$, $\sigma(F)$ was given a value of $|F|/10$.

4.7, ISOTROPIC REFINEMENT

During the least squares refinement of angal, the coordinates of the left handed enantiomorph were retained as it was convenient to calculate a set of phases for this enantiomorph for comparison with phases derived from the symbolic addition procedure.

In the first cycle of the refinement, the nine highest intensity reflections were omitted and all weights were set equal to unity. Atomic positions for the twelve non-hydrogen atoms, an overall temperature factor and six scale factors were refined. Individual scale factors for the six layers of data (l = 0 to l = 5) were refined as the values obtained for these were not very accurate. The R factor was reduced from 0.260 to 0.179. Three cycles of isotropic refinement were then carried out with all the reflections included. The weighting scheme

/described .....

described in section 4.6 was used with W3 = 8. In the second and third of these cycles, isotropic temperature factors of the twelve non-hydrogen atoms were refined. In the next cycle the seventeen highest intensity reflections were omitted and W3 was set equal to 10. A few errors had been discovered in the data and these were corrected. The R factor was reduced to 0.157. The inverse scale factors for the six layers of data were applied to the observed $|F|$ values so that all the data was on a common scale (~10 times absolute). During later stages of the refinement only a single scale factor was refined.

4.8. DIFFERENCE FOURIER SYNTHESIS

A three dimensional difference Fourier synthesis was computed omitting the seventeen highest intensity reflections. It was inspected to see whether there were positive peaks at the calculated hydrogen atom positions (see below). Peaks were found corresponding to the hydrogen atoms on C(1) to C(6). The heights of the peaks were 0.4 to 0.5 $e\text{Å}^{-3}$. The calculated positions of the hydrogen atoms were usually slightly displaced from the peak maxima. Two peaks were found lying along probable hydrogen bond directions. These were interpreted as being the hydrogen atoms attached to O(4) and O(2). It was found later that the position derived for the hydrogen atoms on O(2) gave too long a bond length for the O(2) - H bond (~1.5Å) and a more reasonable estimate of its position was made later though the incorrect position was used during the refinement. The methyl group hydrogen atoms were not placed. The difference

/Fourier ....

Fourier synthesis showed a number of spurious peaks with heights up to 0.6 eÅ^{-3}. There were also some indications of anisotropic vibrations of some of the carbon and oxygen atoms.

4.9, CALCULATION OF HYDROGEN ATOM POSITIONS

The positions of the hydrogen atoms on C(1) to C(5) were calculated on the computer. They were placed such that the C-H bond length was 1.08 Å and the angles between the C-H bond and the other three bonds of the carbon atom, to which the hydrogen atom was attached, were equal. The positions of the hydrogen atoms on C(6) were calculated using a computer program written by Dr. D.A. Rees. The angles between the C-H bonds and the two other bonds of C(6) were assigned values of 109.5 degrees. Again the C-H bond lengths were set equal to 1.08 Å.

4.10, ANISOTROPIC REFINEMENT

Nine of the twelve hydrogen atoms of angal (all except those of the methyl group) were included in the refinement. They were assigned isotropic temperature factors equal to the isotropic temperature factors of the atoms to which they were attached. None of the hydrogen atom parameters were refined. One further cycle of isotropic least squares refinement was carried out varying the isotropic temperature factors and positional parameters of the twelve non-hydrogen atoms and a single scale factor. W3 was set equal to 12. The R factor was reduced to 0.144. It was found that the store of the KDF9 computer was not large enough to enable the simultaneous refinement of the positional parameters and anisotropic temperature factors of the twelve non-hydrogen atoms. The anisotropic temperature factor coefficients β_{ij} were thus refined in separate cycles from

/the ..

the positional parameters. Four cycles of refinement were carried out, β_{ij} values for the twelve non-hydrogen atoms being refined in the first and third cycles and positional parameters in the second and fourth cycles. The final R factor was 0.119. A summary of the least squares refinement based on the film data is given in Table 2. Unobserved reflections were not included in the refinement. The expression used for the anisotropic temperature factor was,

$$\exp - \left(h^2.\beta_{11} + k^2.\beta_{22} + l^2.\beta_{33} + 2.k.l.\beta_{23} + 2.h.l.\beta_{13} + 2.h.k.\beta_{12}\right)$$

4.11, PHASE ANALYSIS

A set of phases, calculated for the refined structure (as the left handed enantiomorph), was used for comparison with the phases derived by the symbolic addition procedure and from the seven atom partial trial structure.

The first set of sixty tangent formula refined phases consisted of 15 two dimensional phases and 45 three dimensional phases. Of the 15 two dimensional phases 14 were correct and 1 was incorrect. For the 45 three dimensional phases the average error in the phase $\Delta\theta$ was 41 degrees. The distribution of the errors is shown in Fig. 11a.

For the expanded set of 136 phases there were 18 two dimensional phases of which 17 were correct. The average error in the three dimensional phases was 52 degrees. The distribution of the errors is shown in Fig. 11b. For the expanded set of 120 phases there were 21 two dimensional phases of which 18 were correct. The average error in the three dimensional phases was 48 degrees. In this case the basic set of 60 phases had been

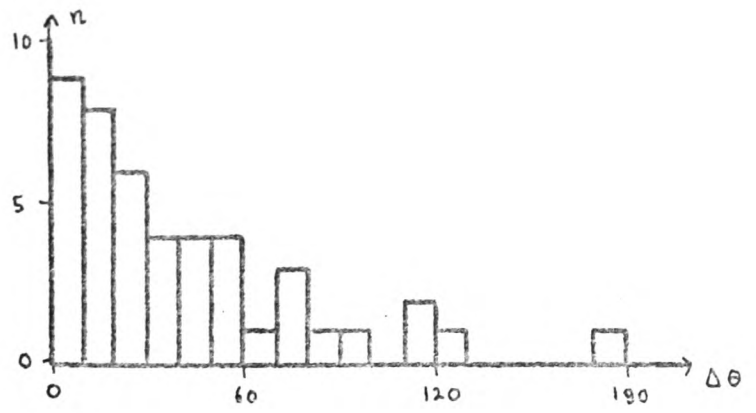
/refined ...

NC	NR	iso/ aniso	R R_1	weights	NA	NP	Parameters varied
0	746	iso	0.260	-	12	0	none
1	737	iso	0.179 325824	unit weights	12	43	12.(x,y,z);6.SF;T_o
2	746	iso	0.168 7809	W3=8	12	43	12.(x,y,z);6.SF;T_o
3	746	iso	0.163 7064	W3=8	12	54	12.(x,y,z,B);6.SF
4	746	iso	0.162 7031	W3=8	12	54	12.(x,y,z,B);6.SF
5	730	iso	0.157 990	W3=10	12	54	12.(x,y,z,B);6.SF
6	730	iso	0.144 556	W3=12	21	49	12.(x,y,z,B);1.SF
7	730	aniso	0.123 393	W3=12	21	73	72.β_{ij};1.SF
8	730	aniso	0.120 379	W3=12	21	37	12.(x,y,z);1.SF
9	730	aniso	0.120 373	W3=12	21	73	72.β_{ij};1.SF
10	730	aniso	0.119 371	W3=12	21	37	12.(x,y,z);1.SF

Notes

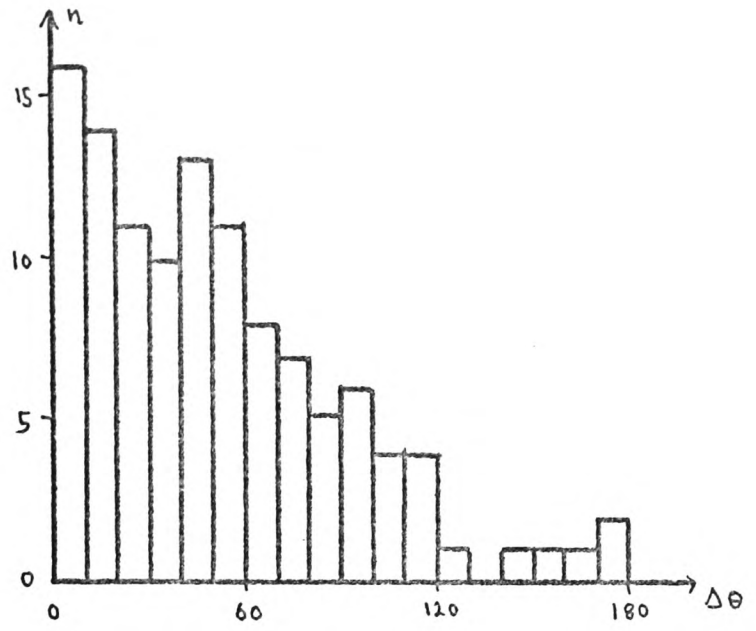
1) NC=number of cycle; NR=number of reflections; NA=number of atoms; NP=number of parameters varied.
2) Nine highest intensity reflections omitted in cycle 1.
3) After cycle 4 some errors in the data were corrected and the 17 highest intensity reflections were omitted.
4) After cycle 5 all the data were placed on a common scale.

Table 2, Summary of least squares refinement of angal based on film data.



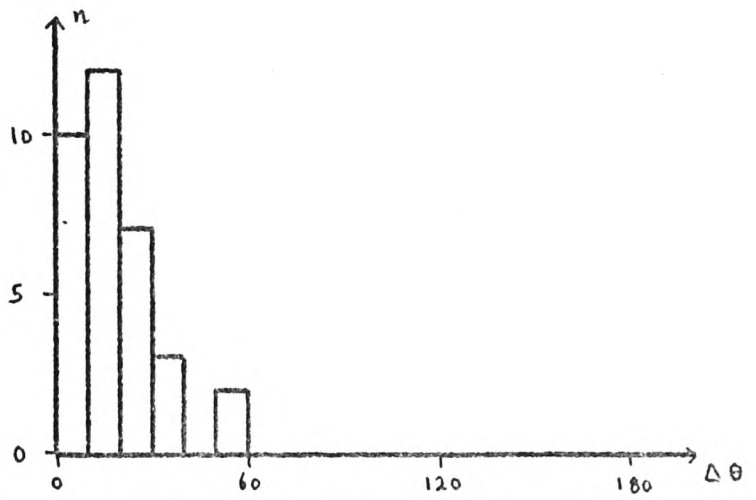
Set 1 of 60 phases.

a.



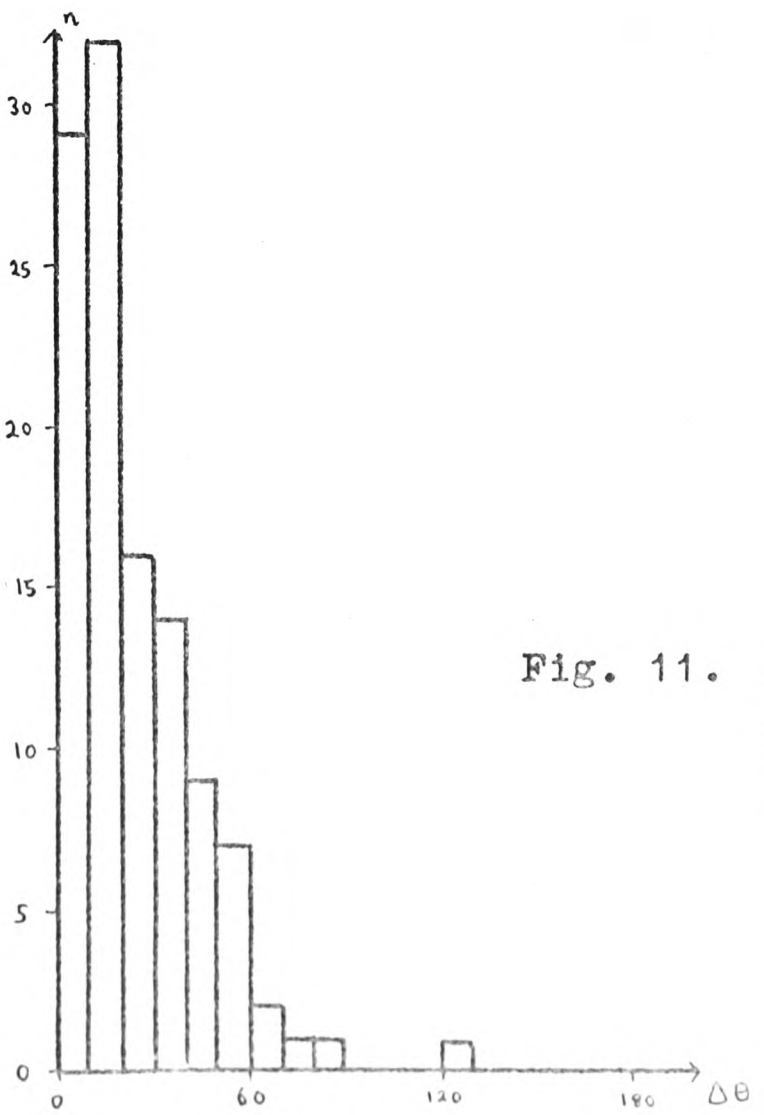
Set of 136 phases.

b.



Phases from 7 atom partial structure. (basic set)

c.



Phases from 7 atom partial structure. (expanded set)

d.

Fig. 11. Distributions of the errors of the three-dimensional phases.

refined along with the additional phases but the average error in the basic set of 60 phases was still 41 degrees and 14 out of the 15 two dimensional phases were correct.

After tangent formula refinement of the basic set of 52 phases derived from the seven atom partial trial structure, 47 phases were accepted. There were 13 two dimensional phases, all of which were correct. The average error in the 34 three dimensional phases was 20 degrees. For the expanded set of 145 phases there were 29 two dimensional phases of which 27 were correct. One of the two incorrect phases, the 104, was the incorrect two dimensional phase in the basic set of 60 phases derived by symbolic addition. The average error in the three dimensional phases was 25 degrees. The distributions of errors for the basic and expanded sets of phases are shown in Fig. 11c and d.

SECTION V: DIFFRACTOMETER DATA

5.1, COLLECTION OF DATA

A second set of intensity data for angal was collected in the department of Biochemistry at the University of Bristol on a Hilger and Watts Y290/FA128 four circle diffractometer controlled by a PDP8 computer. An x-ray tube with a molybdenum target was operated at 40 K.V. and 25 m.A from a Philips 1130 generator. The equipment was kindly made available by Dr. H.C. Watson. Though the author was present at the setting of the crystal, most of the data collection and processing was done in his absence by Dr. H.C. Watson and Dr. P. Wendell.

After the alignment of the diffractometer had been checked, an angal crystal was mounted about its 'c' axis on the goniometer head and was set using the 002 and 040 reflections. One thousand nine hundred and seventy four independent reflections were measured in three ranges of theta, from 0 to 10 degrees, from 10 to 25 degrees and from 25 to 35 degrees. Each reflection was measured both as the hkl and the symmetry related $hk\bar{l}$ reflection. An omega scan was used (19) and the reflections were scanned over 2.4 degrees in sixty steps of 0.04 degrees. The counting time was four seconds for each step, giving a total counting time of eight minutes for each reflection. Between each measurement of the two standard reflections (the 002 and the 040), 48 ordinary reflections were measured. The maximum values of h, k and l were set at 17, 20 and ± 12 respectively.

5.2, PROCESSING OF DATA

The thirty consecutive counts which gave the largest total were taken as representing the intensity of the peak plus the

/background...

background over thirty steps. The other thirty counts were taken as the background over the remainder of the scan. This method of estimating the positions of the reflections in the scan tended to give a systematic over estimation of the intensities of the very weak reflections. If the edge of a reflection came at the end or within one step of the end of the scan, the measurement was rejected. If N_1 was the count for the peak plus background and N_2 was the count for the background, then the intensity of the reflection was taken as being proportional to $N_{pk} = N_1 - N_2$ and the standard deviation based on the counting statistics as $\sigma_{pk} = \sqrt{(N_1 + N_2)}$. The data were corrected for Lorentz and polarisation factors and the two intensity measurements for each reflection were averaged (there was only one measurement for each of the hk0 reflections as $hk0 \equiv hk\bar{0}$).

At two stages in the data collection, after 819 reflections (the end of the 0-25 degree range) and after 1392 reflections, there were slight changes in the measured intensities of the standard reflections. Accordingly, after one cycle of isotropic full matrix least squares refinement based on 819 reflections, the $|F|_{obs}$ data were put on an approximately 10 times absolute scale by applying the following scale factors:-

Reflections nos.	Scale factor
1 - 819	0.3600
820 - 1392	0.3635
1393 - 1974	0.3692

An estimate of the amount of error due to absorption was made by measuring the intensity of a independent reflection as ϕ was rotated in ten degree steps (20). No significant

/systematic ...

systematic variation of intensity was found, though the counting time used for each measurement was rather too short to enable small changes to be detected. The individual intensity measurements varied by up to 3% from the mean value. No absorption correction was applied as the maximum error would not be greater than 2 or 3%.

5.3, WEIGHTING SCHEME

The standard deviations assigned to the $|F|_{obs}$ data were calculated using the expression (21):-

$$\sigma(F) = \frac{1}{2.L_p}\sqrt{\left(\frac{\sigma_{pk}^2 + (0.02\,N_{pk})^2}{N_{pk}}\right)}$$

where L_p = combined Lorentz and polarisation factors.

The factor of 0.02 in the above expression was derived from an inspection of the measurements of the intensity of the 002 reflection (one of the standard reflections) which showed a variation of not more than 2% during the data collection in the theta range of 0 - 25 degrees. The counting errors were small for this reflection and the main source of random error would be due to sources of error such as fluctuation in the x-ray source. For the high intensity reflections this was the significant contribution to the standard deviation whereas for the low intensity reflections the errors based on the counting statistics were important.

/5.4, LEAST ...

5.4, LEAST SQUARES REFINEMENT OF ANGAL USING DIFFRACTOMETER DATA

One cycle of full matrix least squares refinement was carried out using the parameters calculated after the fifth cycle of least squares refinement based on the film data. The weighting scheme incorporated in the computer program was used. A scale factor and the positional parameters and isotropic temperature factors of the twelve non-hydrogen atoms were refined using 819 reflections. The R factor after the cycle was 0.138. A three dimensional difference Fourier synthesis was computed and inspected for hydrogen atom positions. Eight hydrogen atoms showed up, that is all the hydrogen atoms except those of the methyl group and O(2). Nine hydrogen atoms were placed in the same positions as used for the first refinement and each was assigned a temperature factor equal to the isotropic temperature factor of the atom to which it was attached. The three hydrogen atoms of the methyl group were not placed. Structure factors were calculated for the 21 atoms using the full set of data (1974 reflections). The R factor was 0.172.

Four further cycles of full matrix least squares refinement were carried out. The weights were calculated from the supplied $\sigma_{(F)}$ values calculated as described above. In the first and third of these cycles, an overall scale factor and the anisotropic thermal parameters for the twelve non-hydrogen atoms were refined. At no stage were any hydrogen atom parameters refined. One of the hydrogen atoms was slightly misplaced during refinement.

/The ...

The final R factor was 0.097 and the weighted R factor ($= \sqrt{\{ \Sigma \omega_i . ||F|_{obs} - |F|_{calc}|^2 / \Sigma \omega_i |F|^2_{obs}\}}$) was 0.078. An R factor was calculated for the 819 reflections in the 0-25 degree range of theta as this was a comparable set of data to that measured from the films. The R factor was 0.061 compared with 0.119 for the film data (730 reflections). The structure factor tables are given in Appendix II. It can be seen from these that there was a systematic tendency for the weak intensities to be over estimated. The refinement is summarised in Table 3 and two projections of the angal crystal structure are shown in Figs. 12 and 13.

5.5, THE PARAMETERS

The refined parameters* are listed in Tables 4, 5 and 6. Table 4 gives the fractional coordinates with their standard deviations, from the least squares refinement, for the twelve non-hydrogen atoms of angal and the fractional coordinates and isotropic temperature factors calculated for nine hydrogen atoms. Table 5 gives the anisotropic thermal parameters, β_{ij}, for the twelve non-hydrogen atoms and B_{ij} values calculated as suggested by Cruickshank (23). The average value of B_{11}, B_{22} and B_{33} gives an estimate of the mean isotropic temperature factor for the atom. Table 6 lists the bond lengths and bond angles for the non-hydrogen atoms, together with their standard deviations calculated from the standard deviations of the fractional

/coordinates ...

* Note: During the solution and refinement of the angal structure the cell dimensions used were 0.2% less than those quoted in Table 1 and used in the calculation of the inter-atomic distances.

/This ...

NC	NR	iso/ aniso	R R_1	NA	NP	Parameters varied
0	819	-	0.144	12	0	none
1	819	iso	0.138 5736	12	49	12.(x,y,z,B);1.SF
	1974	iso	0.172 111096	21	0	none
2	1974	aniso	0.117 47295	21	73	72.β_{ij};1.SF
3	1974	aniso	0.102 32052	21	37	12.(x,y,z);1.SF
4	1974	aniso	0.098 26625	21	73	72.β_{ij};1.SF
5	1974	aniso	0.097 25952	21	37	12.(x,y,z);1.SF

Notes

1) NC=number of cycle; NR=number of reflections;
NA=number of atoms; NP=number of parameters varied.

2) In cycle 1 the weights were calculated using the routine of the least squares program. In later cycles the weights were calculated as described in section 5.3

Table 3, Summary of least squares refinement of angal based on diffractometer data.

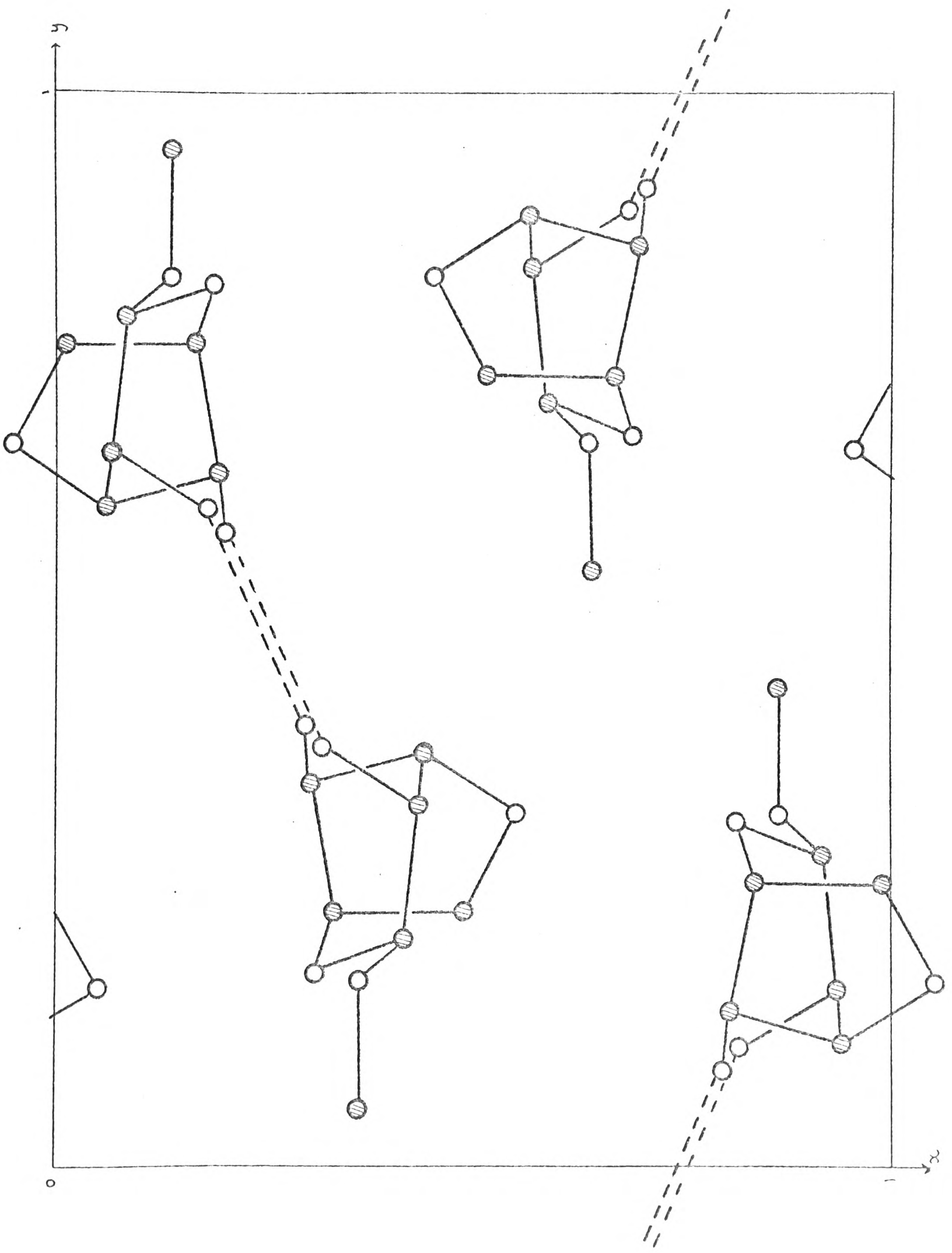


Fig. 12. Projection of the crystal structure of angal (hk0).

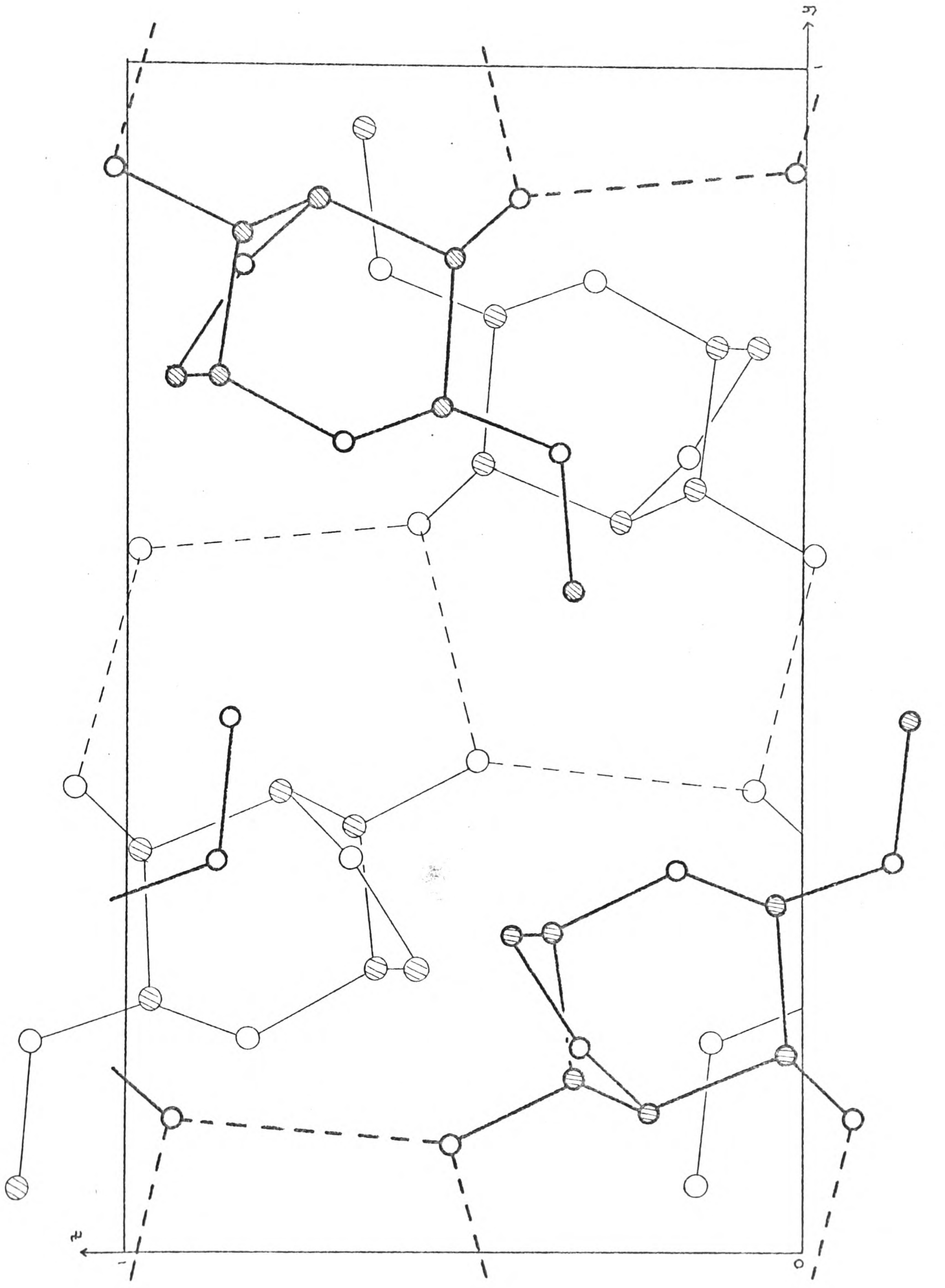


Fig. 13. Projection of the crystal structure of angal (Okl).

Atom	x	y	z
C(1)	0.4159(3)	0.2103(2)	-0.0396(4)
C(2)	0.4347(3)	0.3364(2)	-0.0249(4)
C(3)	0.4411(3)	0.3853(2)	-0.2291(4)
C(4)	0.3042(3)	0.3565(2)	-0.3371(3)
C(5)	0.3344(3)	0.2335(2)	-0.3697(4)
C(6)	0.4880(3)	0.2358(3)	-0.4331(4)
C(7)	0.3635(4)	0.0529(2)	0.1560(5)
O(1)	0.3650(3)	0.1728(2)	0.1342(3)
O(2)	0.3185(3)	0.3879(2)	0.0727(3)
O(3)	0.5502(2)	0.3297(2)	-0.3339(3)
O(4)	0.2973(2)	0.4101(2)	-0.5196(3)
O(5)	0.3145(2)	0.1782(1)	-0.1860(3)

Atom	x	y	z	B
H(1)	0.5073	0.1676	-0.0673	2.73
H(2)	0.5354	0.3553	0.0497	2.87
H(3)	0.4637	0.4733	-0.2256	2.56
H(4)	0.2107	0.3710	-0.2508	2.30
H(5)	0.2678	0.1960	-0.4808	2.23
H(6)	0.4914	0.2520	-0.5909	2.82
H(7)	0.5421	0.1612	-0.3958	2.82
H(8)	0.3250	0.3880	0.2271	3.57
H(9)	0.2750	0.5000	-0.5000	2.73

H(8) on O(2) and H(9) on O(4)

Table 4. Atomic coordinates of angal. Standard deviations in brackets refer to last quoted decimal place of the parameters.

Atom	$10^4 \cdot \beta_{11}$	$10^4 \cdot \beta_{22}$	$10^4 \cdot \beta_{33}$	$10^4 \cdot \beta_{23}$	$10^4 \cdot \beta_{13}$	$10^4 \cdot \beta_{12}$
C(1)	95	43	114	3	-12	-2
C(2)	108	40	107	-10	-19	1
C(3)	98	46	110	-1	-19	-15
C(4)	90	38	87	2	2	7
C(5)	78	43	103	-5	-9	-1
C(6)	100	69	121	-14	4	19
C(7)	189	36	187	23	17	5
O(1)	156	49	107	6	28	12
O(2)	183	52	98	2	7	36
O(3)	70	80	161	3	-7	7
O(4)	127	53	91	7	-11	21
O(5)	110	39	129	7	-19	-11

Atom	B_{11}	B_{22}	B_{33}	B_{23}	B_{13}	B_{12}	B_{ii}
C(1)	3.39	2.49	2.18	0.10	-0.31	-0.09	2.69
C(2)	3.85	2.31	2.05	-0.33	-0.50	0.05	2.74
C(3)	3.49	2.66	2.11	-0.03	-0.50	-0.68	2.75
C(4)	3.21	2.20	1.67	0.07	0.05	0.32	2.36
C(5)	2.78	2.49	1.97	-0.17	-0.24	-0.05	2.41
C(6)	3.56	3.99	2.32	-0.47	0.10	0.86	3.29
C(7)	6.73	2.08	3.58	0.77	0.44	0.02	4.13
O(1)	5.56	2.83	2.05	0.20	0.73	0.54	3.48
O(2)	6.52	3.01	1.88	0.07	0.18	1.63	3.80
O(3)	2.50	4.62	3.08	0.10	-0.18	0.32	3.40
O(4)	4.53	3.07	1.74	0.23	-0.29	0.95	3.11
O(5)	3.92	2.26	2.47	0.23	-0.50	-0.50	2.88

Table 5, Anisotropic thermal parameters of angal.

Bond	Length(A) (Diff. data)	Length(A) (Film data)
C(1)-C(2)	1.533(3)	1.575(10)
C(2)-C(3)	1.534(4)	1.521(13)
C(3)-C(4)	1.535(4)	1.532(12)
C(4)-C(5)	1.526(3)	1.549(11)
C(5)-C(6)	1.518(4)	1.506(12)
C(1)-O(5)	1.449(3)	1.442(11)
C(1)-O(1)	1.374(4)	1.350(11)
C(2)-O(2)	1.432(4)	1.449(11)
C(3)-O(3)	1.429(3)	1.413(11)
C(4)-O(4)	1.422(3)	1.416(10)
C(5)-O(5)	1.449(3)	1.426(12)
C(6)-O(3)	1.449(4)	1.470(11)
C(7)-O(1)	1.453(3)	1.445(10)

Angle	(Diff. data)	(Film data)	(Angles in degrees)
C(2)-C(1)-O(5)	112.8(2)	112.5(6)	
C(2)-C(1)-O(1)	108.0(2)	107.9(6)	
O(5)-C(1)-O(1)	107.1(2)	108.9(6)	
C(1)-C(2)-C(3)	108.9(2)	108.4(6)	
C(1)-C(2)-O(2)	111.9(2)	109.4(6)	
C(3)-C(2)-O(2)	107.4(2)	107.7(6)	
C(2)-C(3)-C(4)	109.3(2)	110.2(6)	
C(2)-C(3)-O(3)	108.5(2)	108.7(6)	
C(4)-C(3)-O(3)	104.8(2)	105.6(6)	
C(3)-C(4)-C(5)	97.7(2)	97.5(6)	
C(3)-C(4)-O(4)	111.7(2)	111.6(6)	
C(5)-C(4)-O(4)	108.5(2)	109.1(6)	
C(4)-C(5)-O(5)	107.0(2)	106.8(6)	
C(4)-C(5)-C(6)	101.8(2)	100.5(6)	
O(5)-C(5)-C(6)	112.8(2)	113.6(6)	
C(5)-C(6)-O(3)	105.4(2)	105.9(6)	
C(7)-O(1)-C(1)	114.9(2)	115.8(6)	
C(3)-O(3)-C(6)	108.3(2)	107.6(6)	
C(1)-O(5)-C(5)	114.0(2)	116.2(6)	

Table 6. Bond lengths and bond angles of angal. Standard deviations in brackets refer to the last quoted decimal place of the parameters.

coordinates as described by Ahmed and Cruickshank (24). The bond lenths and bond angles derived from the film data are included for comparison.

Note cont'd:

This meahs that there will be some slight, though probably insignificant, errors in the thermal parameters, $|F|_{calc}$ values and hydrogen atom positions. For work requiring accurate hydrogen atom positions the hydrogen atoms positions should probably be recalculated as, in addition to any small error due to the different cell dimensions, the positions of the atoms, to which the hydrogen atoms were attached, shifted slightly during the refinement.

SECTION VI: DISCUSSION OF THE STRUCTURE

6.1, DISCUSSION OF THE ANGAL STRUCTURE

In estimating a set of coordinates for angal (see Part II, section 3.9) the structure was treated as if it were a five membered sugar ring (-C(3)-C(4)-C(5)-C(6)-O(3)-) fused on to part of a six membered pyranose ring (-C(5)-O(5)-C(1)-C(2)-C(3)-). The geometry of the angal molecule as determined by x-ray diffraction confirms the reasonableness of this treatment. In almost every part of the molecule the parameters are consistent with the general trends for pyranose rings and for five membered rings as exemplified by furanose rings. Though the five membered ring of angal is not a furanose ring, furanose rings were the closest analogues for which there was a reasonable quantity of crystal data for comparision.

The information about bond lengths and bond angles in sugar rings was obtained from a review by Jeffrey and Rosenstein (25), from a survey of the crystal structures of seventeen compounds containing pyranose rings and four compounds containing furanose rings and from summaries of the general trends given by Chu and Jeffrey (26), Berman, Chu and Jeffrey (27) and Sundaralingam(28).

There appears to be no significant difference in C-C and C-O bond lengths (excepting C(1)-O(1)) between pyranose and furanose rings. The average C-C bond length in angal of 1.529Å is very close to the average value of 1.527Å for pyranose rings. The average C-O bond length of 1.441Å is only marginally greater than the average value for pyranose rings of 1.430Å.

The range of bond angles at the pyranose ring carbon atoms (C(1) and C(2)) is typical of that found for pyranose rings with

/an ....

an average value of 109.3 degrees very close to the tetrahedral angle of 109.5 degrees. The O(5) ring oxygen angle of 114.0 degrees is very close to the average of 113.9 degrees. The bond angle at O(1) of 114.9 degrees is only slightly greater than the values of 113.2, 113.4 and 113.2 degrees found for the methyl glycosides of α-D-galactose-6-bromohydrin (29), β-xylose(30) and β-maltose (26).

A systematic shortening of the C(1)-O(1) bond has been observed in most pyranose sugars, exceptions being α-methyl-galactoside-6-bromohydrin (29) and sucrose (31), both of which have O(1) axial and the hydrogen atom on O(1) substituted. Chu and Jeffrey (26) have suggested the following empirical 'rules' for the occurrence of C(1)-O(1) bond shortening:-

(i) An axial glycosidic bond is shortened only if the hydrogen atom attached to O(1) is unsubstituted.

(ii) An equatorial glycosidic bond is shortened irrespective of whether the hydrogen atom attached to O(1) is substituted or not.

(iii) Where the C(1)-O(1) bond is shortened, no significant difference occurs in the C-O (ring) bond lengths.

(iv) Where there is no shortening of the C(1)-O(1) bond, the C-O(ring) bond lengths are possibly different, the bond adjacant to the glycosidic bond being the shorter.

According to these rules, shortening of the C(1)-O(1) bond should be observed in angal and there should be no significant asymetry in the lengths of the C-O(ring) bonds. The

/structure ...

structure confirms these predictions. The C(1)-O(1) bond has a length of 1.374 Å which is 0.067 Å (>18∝) shorter than the average C-O bond length. The C-O(ring) bonds are 1.449 Å, C(1)-O(5), and 1.450 Å, C(5)-O(5), showing no significant asymmetry. The C(1)-O(1) bond length is slightly shorter than the average length for this bond of 1.390 Å.

In furanose rings the average C-C-O bond angle is 105.8 degrees. The two C-C-O bond angles of angal at C(3) and C(6) have values close to this (104.8 and 105.4 degrees). The C-C-C bond angles within the ring tend to be smaller than the C-C-O angles with an average value of 102.5 degrees at the carbon atoms lying in the four atom plane of the ring and 100.9 degrees at the carbon atom lying out of the four atom plane. The observed values for angal were 101.8 degrees (at C(5)) and 97.7 degrees (at C(4)), both following the general trend, though the angle at C(4) is rather smaller than average. The ring oxygen bond angle of 108.3 degrees is close to the average of 109.8 degrees.

The other angles at C(3) and C(5) lie in the range 107.0 to 112.8 degrees with an average value of 109.4 degrees close to the tetrahedral angle.

The distance across the pyranose ring from C(3) to C(5) is 2.30 Å which 0.2 Å less than the value of 2.50 Å for an unbridged pyranose ring. The best plane through the four atoms C(2), C(3), C(5) and O(5) was calculated. The distances of the four atoms from this plane were 0.015, 0.017, 0.018 and 0.012 Å

/respectively ..

respectively. C(1) was 0.543 Å out of the plane and C(4) was 0.921 Å out of the plane. The average distance out of the plane for C(1) and C(4) for a glucopyranose residue is about 0.66 Å(32).C(4) of angal is forced further out of the plane by the formation of the anhydro bridge. The pyranose ring on the other hand is slightly flattened at C(1). This flattening could relieve two sources of strain. First, tendency for O(5) and C(2) to be pushed apart when the anhydro ring is formed (the distance between C(2) and O(5) is 2.48 Å compared with 2.40 Å for a glucopyranose residue) and second, to relieve a rather short contact distance between H(1) on C(1) and H(7) on C(6). If there was no flattening of the ring the distance would be about 1.8-1.9 Å which is slightly less than the suggested 'outer limit' H- - -H contact distance of 1.9 Å (32). The distance between the calculated hydrogen atom positions on angal was 2.3 Å.

The best plane through C(3), O(3), C(6) and C(5) of the five membered ring was also calculated, the distances of the four atoms from the plane being 0.015, 0.018, 0.023 and 0.014 Å respectively. C(4) was 0.706 Å out of the plane.

Two projections of the molecular structure of angal are shown in Fig. 14.

The position of the methyl group was also of interest. The conformational energy based on non-bonded atomic interactions was calculated as a function of the angle of rotation around the C(1)-O(1) bond (Fig. 15). The Kitaygorodsky potential energy

/function ....

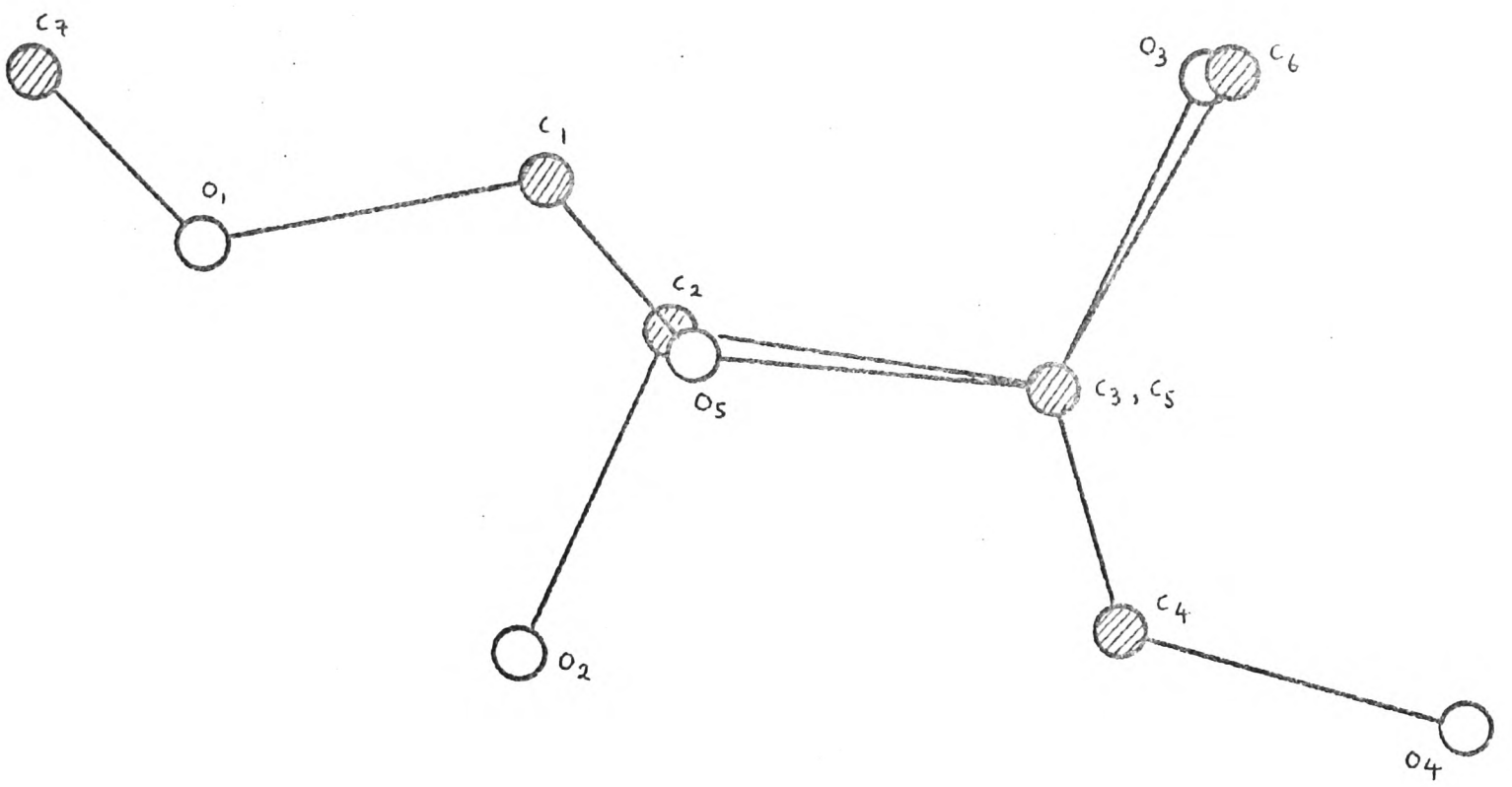


Fig. 14. Two molecular projections of angal.

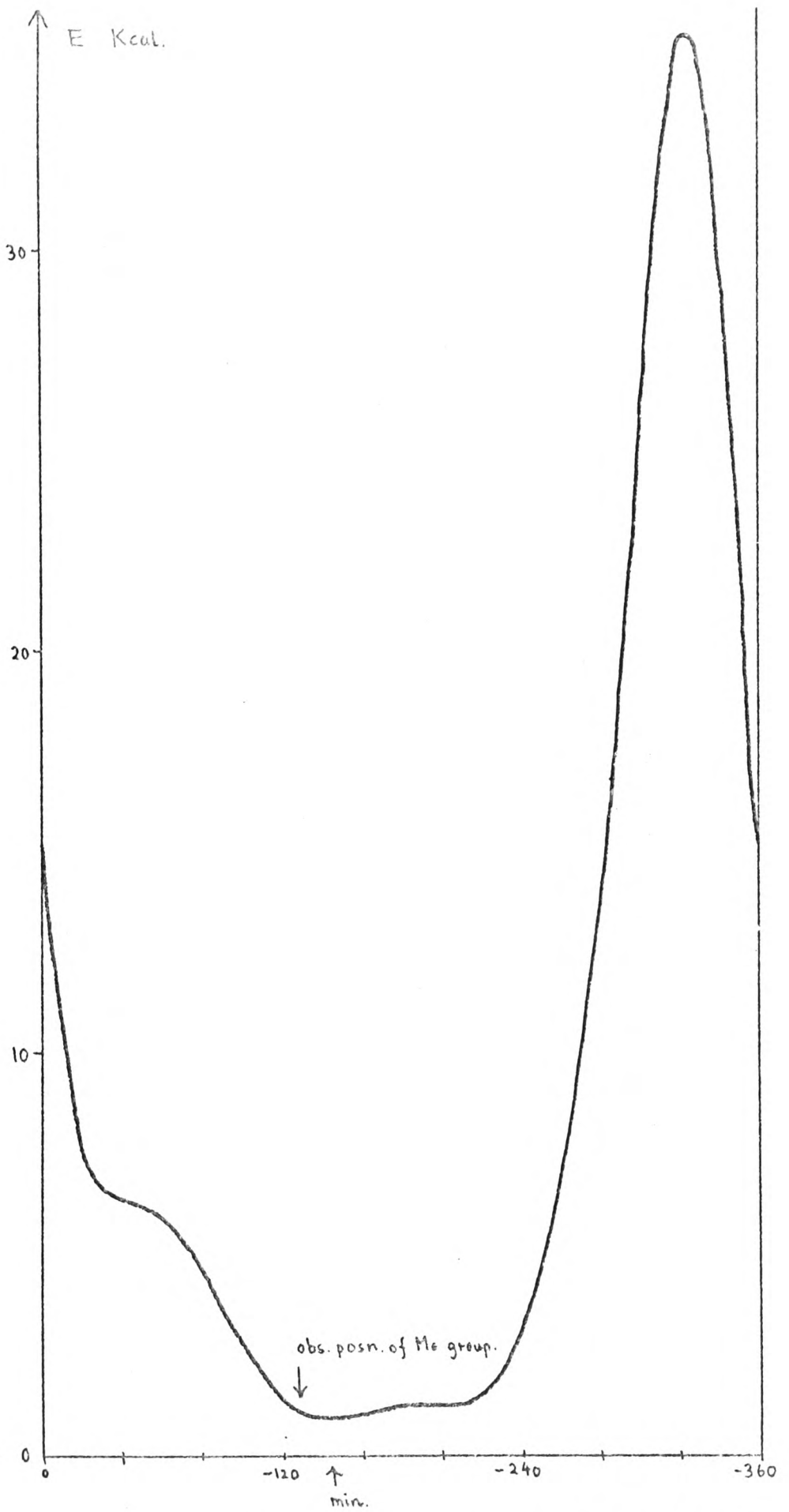


Fig. 15. Potential energy profile for methyl group of angal.

function (67) was used with r_o values of 3.92, 3.76 and 3.33 Å for the C - - -Me, O - - -Me and H - - -Me interactions (69). The calculation was carried out using the coordinates of the calculated angal model (see Part II, Section 3.9). The methyl group appeared to have a fair degree of freedom as shown by the broad potential energy minimum. The observed dihedral angle C(7)-O(1)-C(1)-C(4) of -125.5 degrees was very close to the calculated potential energy minimum position. The position of the methyl group may also be explained in terms of dipole interactions between the unshared electron pairs on O(1) and O(5) (33) and is consistent with the position of the methyl group in other similar monosaccharides. The dihedral angle O(5)-C(1)-O(1)-C(7) is 68.1 degrees for angal compared with 72.3, 69.2 and 64.8 degrees for the methyl glycosides of β-xylose(30), β-maltose(26) and 1-thio-β-D-xylose (34). C(7) shows quite large anisotropic vibrations in the plane approximately normal to the O(1)-C(7) bond.

Two hydrogen bonds were found involving O(2) and O(4) each of which is hydrogen bonded to the other in two different molecules. As in the case of other methyl glycosides (61) no hydrogen bonds involving the ring oxygen atom O(5) were found. The lengths of the two hydrogen bonds were 2.845 Å and 2.744 Å and the angal molecules are hydrogen bonded to give double chains of molecules running parallel to the 'c' axis as shown in Fig. 16. These double chains would appear to vibrate fairly strongly in a direction approximately parallel to the 'a' axis as shown by the fact that the average B_{11} value is 4.17 compared with 2.83 for B_{22} and 2.26 for B_{33}.

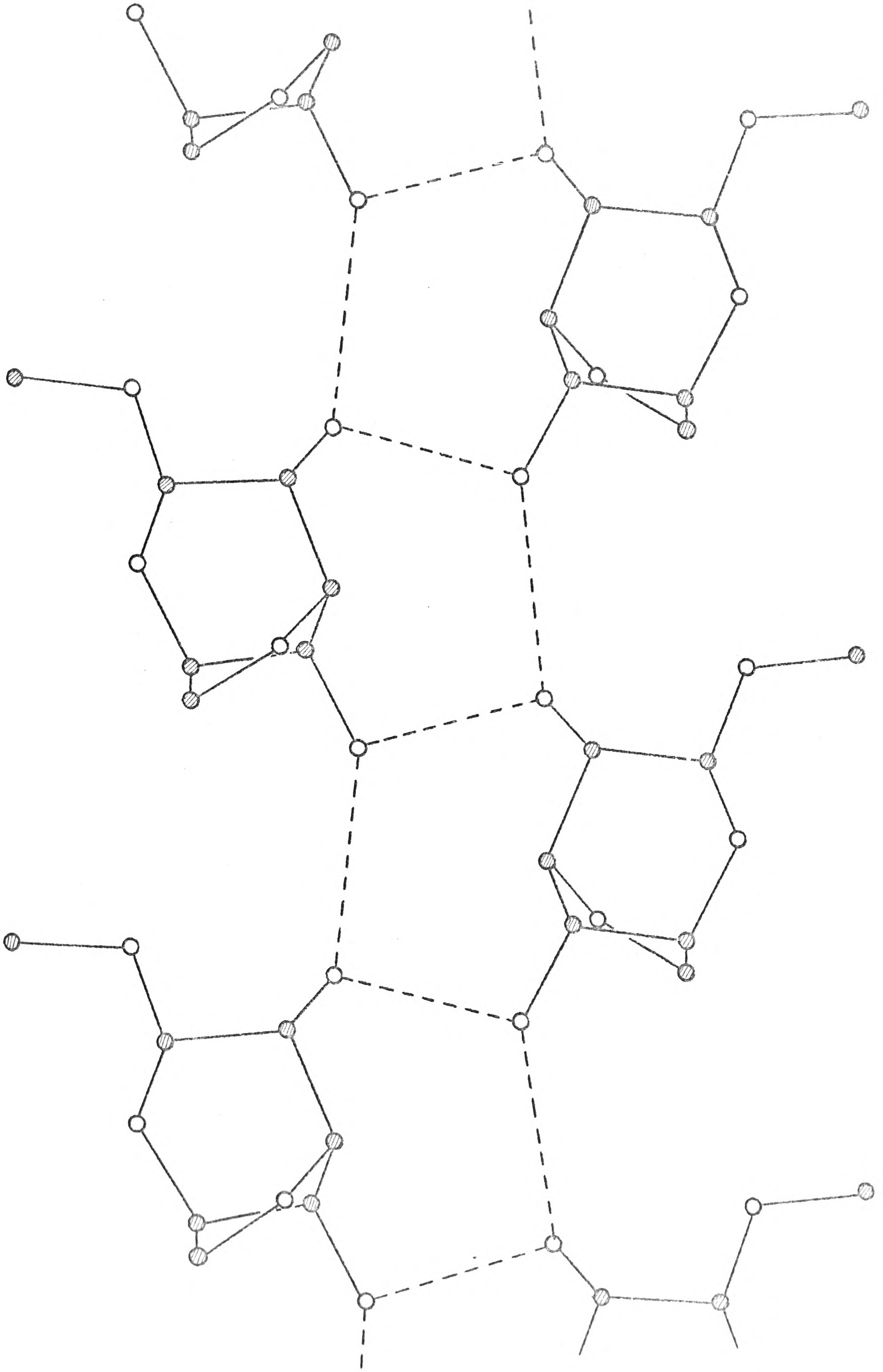


Fig. 16. Hydrogen bonded double chains of angal.

SECTION VII: OTHER METHODS TRIED

7.1, SYSTEMATIC ROTATION OF ANGAL

In the early stages of the structure determination an attempt was made to solve the angal structure in projection (hk0) by rotating a model of the molecule systematically about the ring centre position as determined from the two dimensional Patterson function. For each orientation of the molecule a set of structure factors and an agreement factor were calculated. The orientation of the molecule was defined by three variable Eulerian angles θ_1, θ_2 and θ_3 where θ_1 was a rotation about the 'Z' axis, θ_2 a rotation about the new 'x' axis and θ_3 a rotation about the new 'Z' axis. The rotation matrix was,

$$M = \begin{vmatrix} -\sin\theta_1 \cdot \cos\theta_2 \cdot \sin\theta_3 + \cos\theta_1 \cdot \cos\theta_3 & \cos\theta_1 \cdot \cos\theta_2 \cdot \sin\theta_3 + \sin\theta_1 \cdot \cos\theta_3 & \sin\theta_2 \cdot \sin\theta_3 \\ -\sin\theta_1 \cdot \cos\theta_2 \cdot \cos\theta_3 - \cos\theta_1 \cdot \sin\theta_3 & \cos\theta_1 \cdot \cos\theta_2 \cdot \cos\theta_3 - \sin\theta_1 \cdot \sin\theta_3 & \sin\theta_2 \cdot \cos\theta_3 \\ \sin\theta_1 \cdot \sin\theta_2 & -\cos\theta_1 \cdot \sin\theta_2 & \cos\theta_2 \end{vmatrix}$$

A computer program to carry out the calculation was written by Dr. M.M. Harding. To save computing time, the observed $|F|$ values were divided by the square root of the local average $|F^2|_{obs}$ values to give reduced structure factors $F_r(hkl)$. The expression used for calculating these reduced structure factors was

$$F_r(hkl)_{calc} = \sum_j^N . Z_j . \exp\left[-2.\pi.i\left(h.x_j + k.y_j + \ell.z_j\right)\right]$$

where Z_j = atomic number of the jth atom in a unit cell of N atoms.

The simplified form of the geometric structure factor was

/used:- ...

used:-

$$F_r(hk0) = 4 \sum_i Z_i \cdot \cos(2.\pi.h.x_i).\cos(2.\pi.k.y_i) \quad \text{for } h+k \text{ even}$$

$$F_r(hk0) = -4 \sum_i Z_i \cdot \sin(2.\pi.h.x_i).\sin(2.\pi.k.y_i) \quad \text{for } h+k \text{ odd}.$$

A table of sine values was calculated and stored in order to save computing time.

The calculation of the set of coordinates for angal is described in Part II, section 3.9. Eleven of the twelve non-hydrogen atoms were included in the calculation, the methyl group carbon atom being omitted. Intervals of 15 degrees were chosen for θ_1, θ_2 and θ_3 with ranges of 0-360, 0-180 and 0-360 degrees respectively. Structure factors were calculated for hk0 reflections with h from 0 to 6 and k from 0 to 6. It was desirable to use low order reflections as the structure factors were less sensitive to atomic positions. About 50 seemed a reasonable number of reflections to make comparison between the observed and calculated structure factors worth while. The calculated structure factors were scaled such that

$$\sum |F_r|_{obs} = \sum |F_r|_{calc}$$

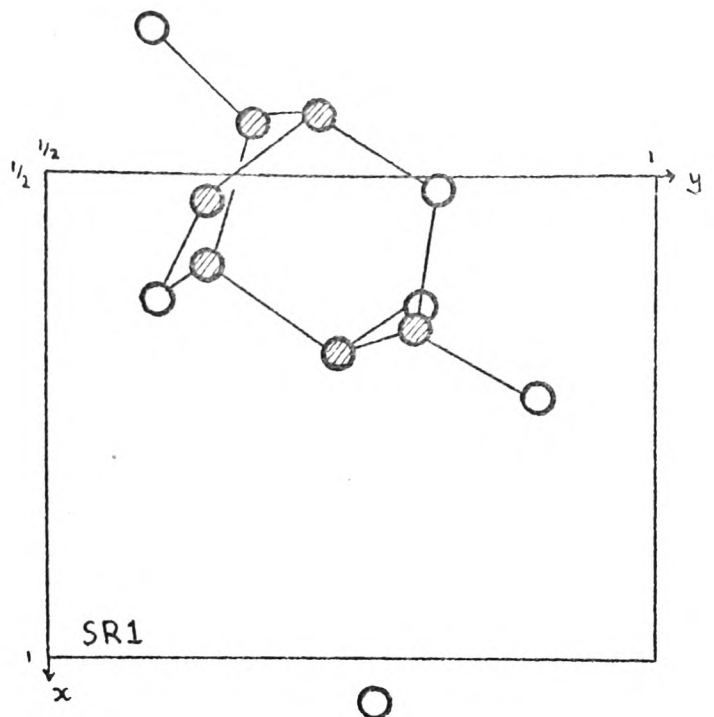
and two agreement factors were calculated

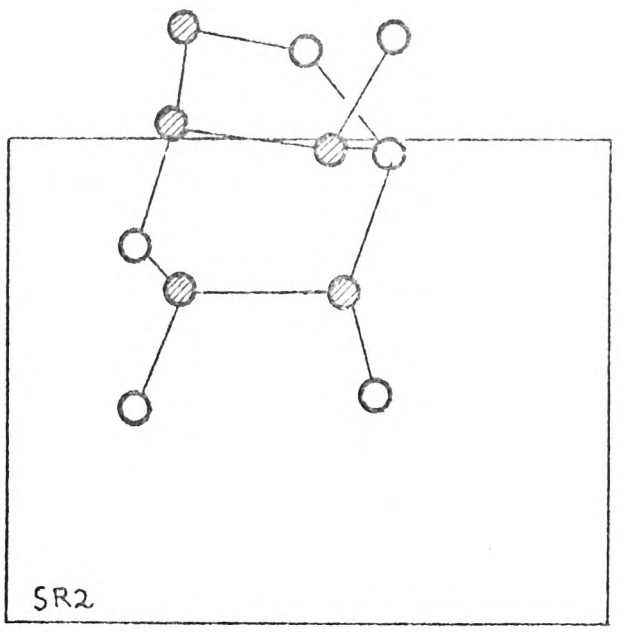
$$R_1 = \sum | |F_r|_{obs} - |F_r|_{calc} |$$

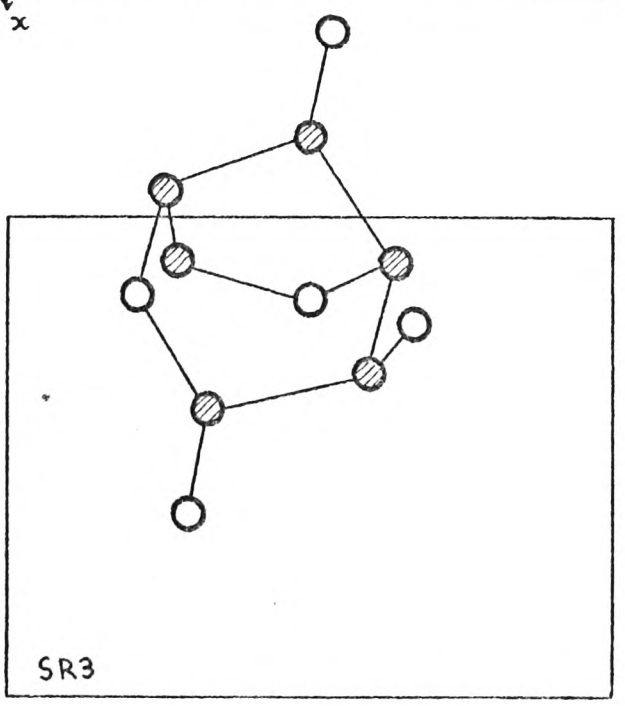
$$R_2 = \sum (|F_r|_{obs} - |F_r|_{calc})^2$$

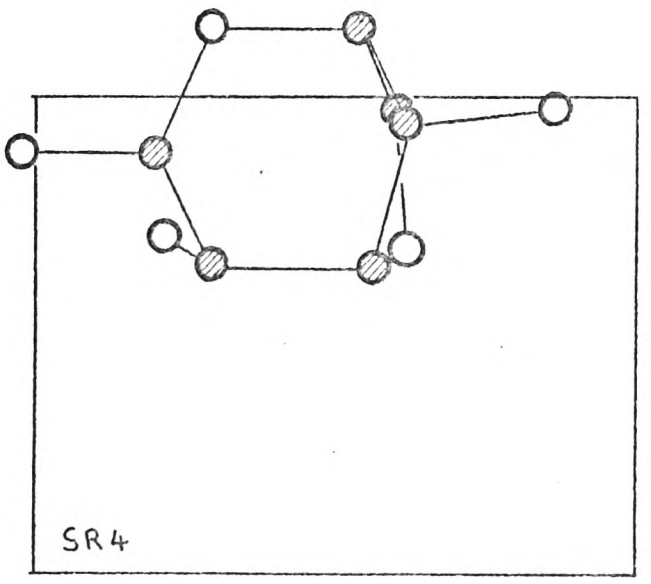
The agreement factor R_2 was plotted out as a function of θ_1, θ_2 and θ_3 and the plots were inspected for minimum values. Several minima were found and the structures corresponding to the five lowest minima, SR1-SR5, are shown in Fig. 17. None of these

/structures ...









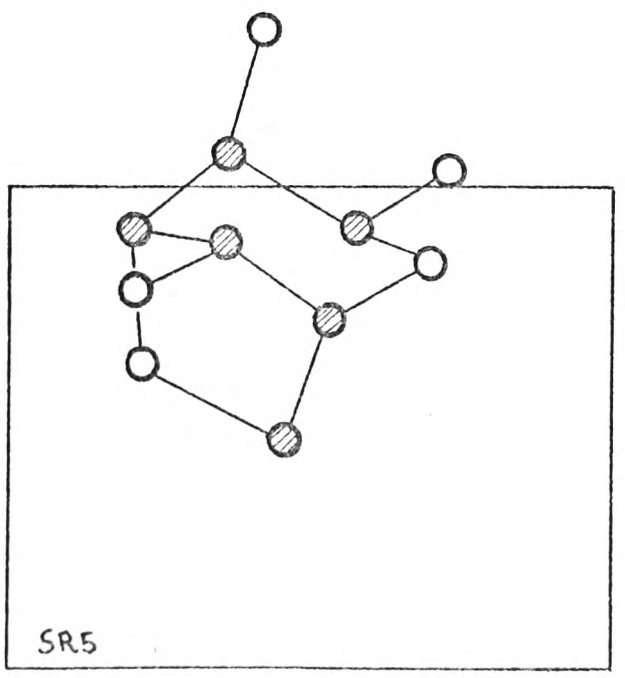


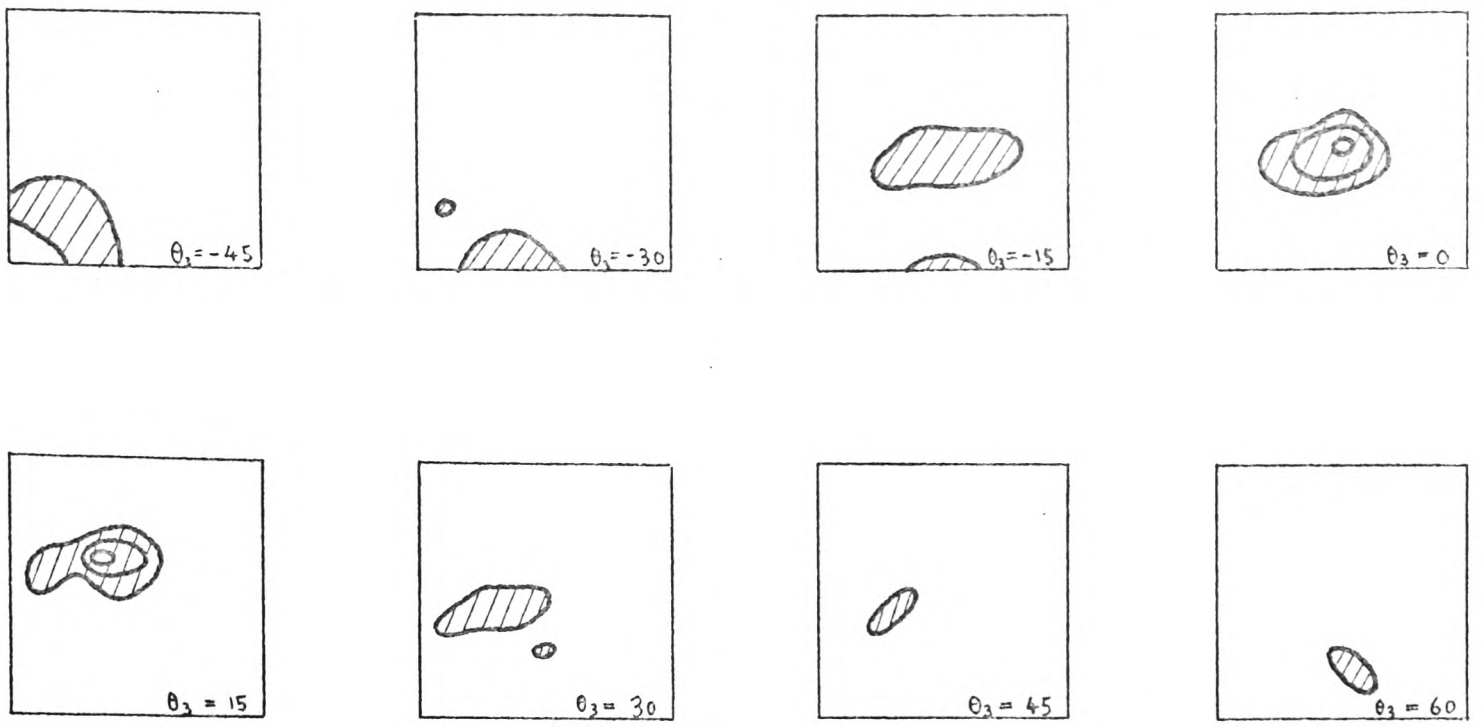
Fig. 17. The best structures from the systematic rotation of angal.

structures refined satisfactorily though SR1 gave fairly good agreement between observed and calculated structure factors at low $\sin\theta$.

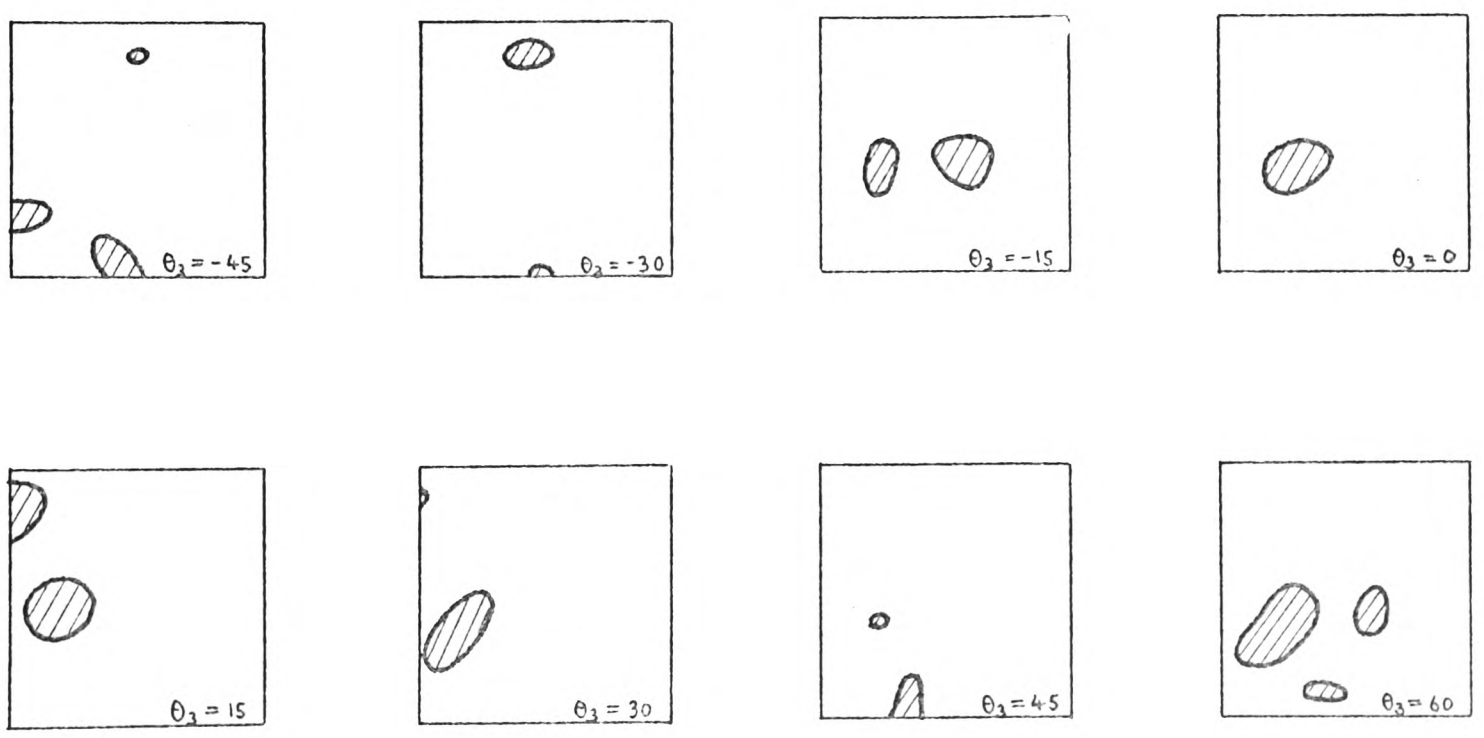
From the correct structure, as derived by direct methods, it was apparent that the ring centre position determined from the two dimensional Patterson function was misplaced by about 0.5 Å in the x coordinate and 0.2 Å in the y coordinate. There were also errors in the measurement of some of the high intensity reflections. Part of the calculation was therefore repeated with corrected data, covering the range in which the correct orientation was to be found. The calculation was done both for the correct ring centre position and the ring centre position calculated from the Patterson function.

The results are shown in Fig. 18. For the rotation around the correct ring centre position there is a good minimum which gives an orientation of the molecule very close to that in the correct structure (Fig. 19), whereas at the Patterson function ring centre-position there is a minimum at approximately the same orientation but it is not significantly better than several other minima round about it. Probably the misplaced ring centre position was a more important factor than the few rather poorly estimated intensities in causing the systematic rotation method to fail for angal. The good minimum obtained in the repeated calculation at the correct ring centre position indicates that the model used was satisfactory and that the omission of the methyl group was unimportant.

/7.2, ROTATION ...



a. (correct centre.)



b. (Patterson function centre.)

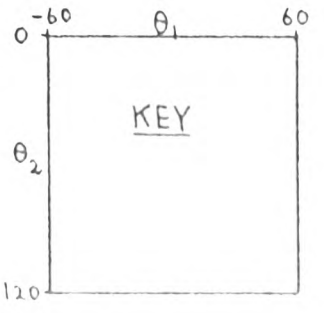
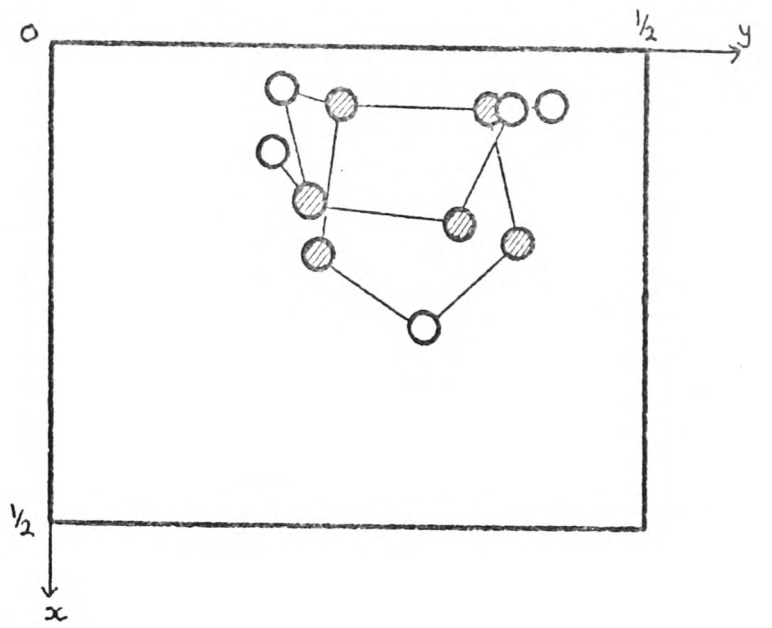


Fig. 18. The positions of the minima from the re-run of part of the systematic rotation of angal.
(θ_1, -60° to +60°; θ_2, 0° to +120°; θ_3, -45° to +60°)



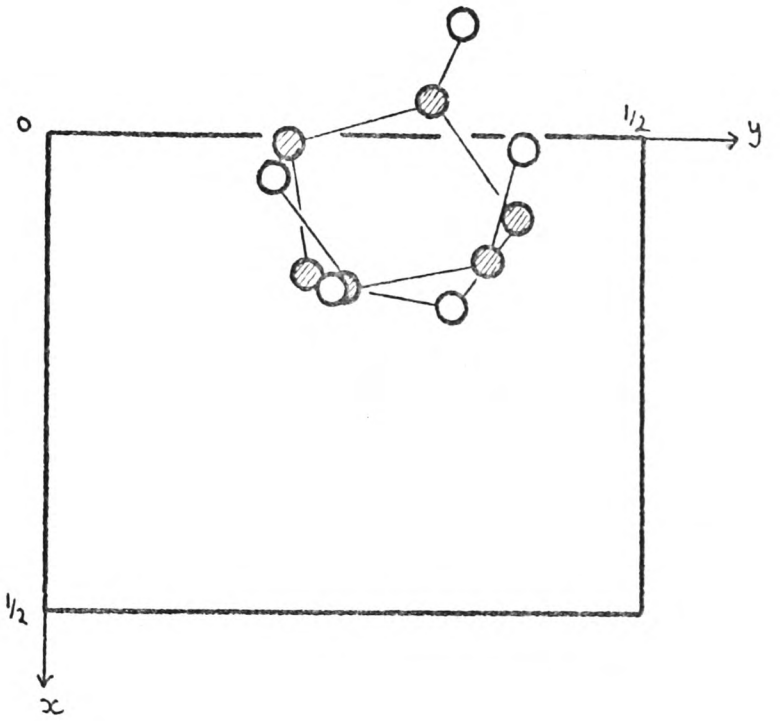


Fig. 19. The best orientations of angal at (a) the correct ring centre position and (b) the ring centre position derived from the two dimensional Patterson function. (From the re-run of part of the systematic rotation.)

7.2, ROTATION FUNCTIONS (35,21)

Various rotation functions were calculated in order to try and find the orientation of the angal molecule which best fitted the three dimensional Patterson function. The method involves taking a vector model of the molecule or a known fragment of the molecule and rotating it around the origin of the three dimensional Patterson function to find the position of 'best fit' (35).

The rotation functions were calculated on the computer using a vector model calculated from the estimated set of angal coordinates (Part II, section 3.9). The maximum length of vector considered was 3.9 Å and the three dimensional Patterson function was stored on a grid of approximately 0.25 Å spacing. Only the asymmetric part of the Patterson function was stored. The value of the Patterson function at the end of each vector was taken as being that of the nearest grid point. The orientation of the vector model was defined by three variable Eulerian angles as for the systematic rotation described above. Ten degree intervals were taken in the three angles. The ranges of the angles required depend on the symmetry of the vector model and of the Patterson function (21). Though the vector model of a single molecule of angal had symmetry $\bar{1}$, the presence of the four orientations of the molecule in the unit cell enabled it to be treated as mmm. The symmetry of the Patterson function was also mmm and thus the ranges required for θ_1, θ_2 and θ_3 were 0-90, 0-90 and 0-180 degrees respectively. Various criteria can be used to determine the 'best fit' of the vector model to the Patterson

/function ...

function. Two of these are the minimum function and the sum function (21,36).

The rotation minimum function is defined by

$$M_{\theta_1\theta_2\theta_3} = \min\left(P_1, P_2, P_3 \cdots P_q\right)$$

where P_i is the value of the Patterson function at the end of the ith vector and q is the total number of vectors in the model.

If this function is positive then the ends of all the vectors lie in positive regions of the Patterson function. The value of the minimum function gives the value of the Patterson function at the end of the 'worst' fitting vector, that is the one which falls in the least positive region. The function can be modified to take into account the weights of the different vectors.

In this form the function is

$$M_{\theta_1\theta_2\theta_3} = \min\left(\frac{1}{Z_1}.P_1, \frac{1}{Z_2}.P_2 \cdots \frac{1}{Z_q}.P_q\right)$$

where Z_i is the weight of the ith vector.

The minimum function is a very sensitive test. If it is to be successful the single weight vectors must be resolved in the Patterson function. The vector model must be accurate and fine intervals of θ_1, θ_2 and θ_3 are required. A weighted minimum function was calculated for angal with the weights of the vectors being taken as the product of the atomic numbers of the atoms at the end of the vectors. No positive regions were found and so it was inspected for the least negative regions.

/A ...

UNIV. EDINBURGH EX BIBL

A modified form of the minimum function was also calculated. In this function a search was carried out over the 27 grid points surrounding the nearest grid point to the end of each vector and the maximum value of the Patterson function taken.

$$M(JWC)_{\theta_1\theta_2\theta_3} = \min\left(\frac{1}{Z_1} \cdot \max(P_1) , \frac{1}{Z_2} \cdot \max(P_2) \cdots \frac{1}{Z_q} \cdot \max(P_q) \right)$$

where max (P_i) = maximum value of the Patterson function at the 27 grid points surrounding the end of the ith vector.

The M(JWC) function was designed to use the minimum function criterion of best fit but to lower the sensitivity of the method thereby allowing for slight errors in the atomic positions of the model or for the use of too large angular intervals in θ_1, θ_2 and θ_3. The function has the disadvantage that a certain amount of distortion is introduced which could possibly lead to a false solution though the correct solution should also give a maximum. The calculation time is greater than that of the minimum function for comparable angular ranges. The search over the 27 grid points need not however be carried out each time for each vector. If the value of the Patterson function at the nearest grid point to the end of a vector is greater than the minimum value found for $\frac{1}{Z_i} \cdot \max(P_i)$ at that stage then no search need be carried out for that vector.

In the M(JWC) function calculated for angal there were a few small positive peaks but no definite indication of the best orientation of the molecule. The results from the minimum functions suggested that either the model was in error or that

/the ....

the single weight vectors were not resolved in the Patterson function. From the final structure it was found that a few of the single weight vectors came in quite negative regions of the Patterson function, though usually fairly near high positive peaks. This could have been caused by diffraction effects round the high peaks.

A second type of rotation function is the rotation sum function:-

$$S_{\theta_1 \theta_2 \theta_3} = \sum_{i=1}^{q} P_i$$

The value of the sum function should be large and positive when the ends of all the vectors lie in positive regions of the Patterson function. There may however be spurious maxima caused by the chance coincidence of even a low weight vector with a high peak in the Patterson function. If, however, a few atoms of the model are slightly uncertain in their positions then the sum function is preferable to the minimum function.
sum function was calculated for angal. The five best orientations
A rotation obtained from it are shown in Fig. 20. The ring orientations in D and E were not consistent with the ring orientation found from the visual inspection of the Patterson function. The orientation of the molecule in B was however very close to that in the correct structure.

7.3, SYSTEMATIC TRANSLATION

Having determined the possible orientation of a molecule, it is necessary to find its position in the unit cell. This

/may ...

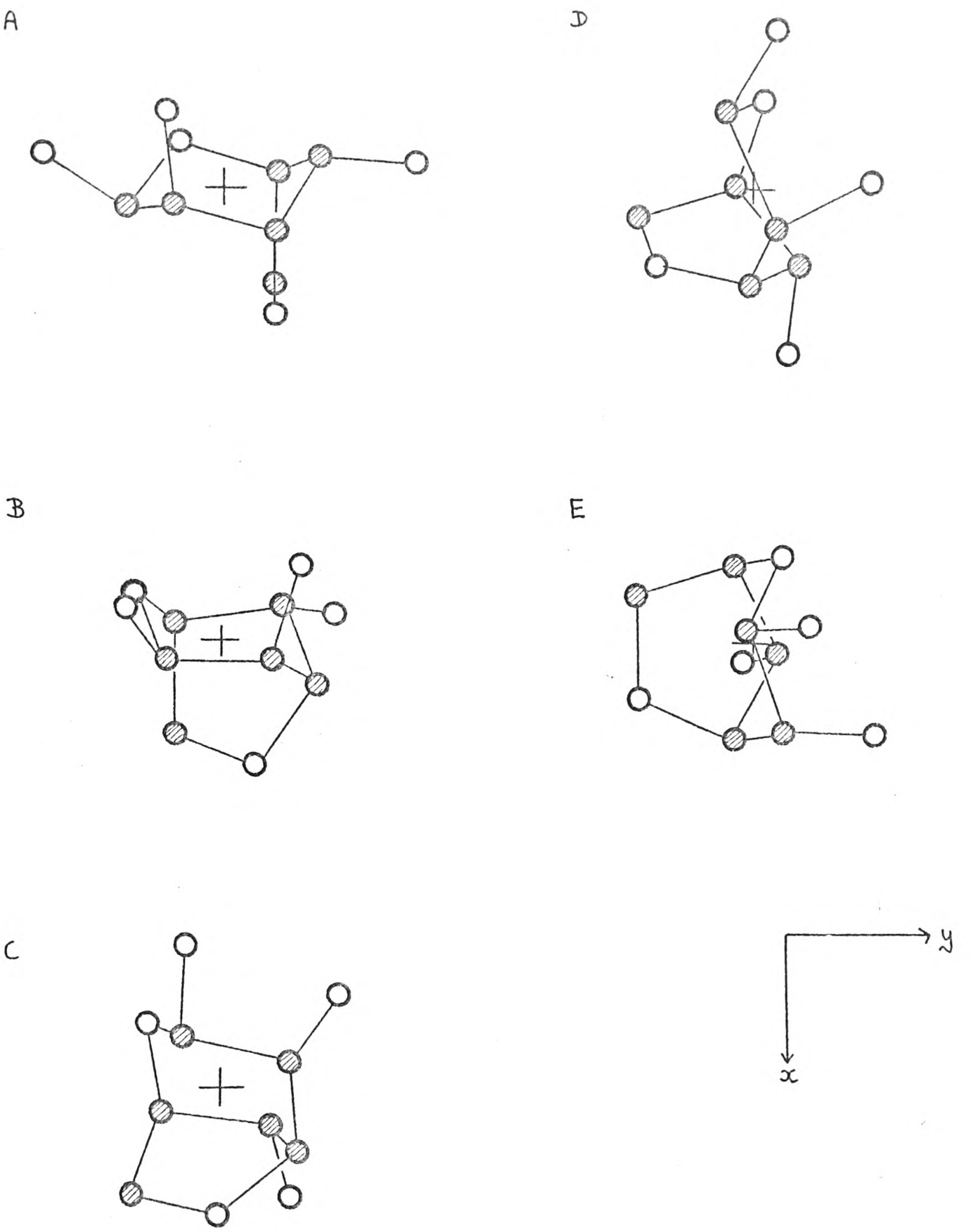


Fig. 20. The five best orientations of angal from the rotation sum function.

may be done by translating the molecule, in the proposed orientation, systematically and stepwise over the unit cell and calculating sets of structure factors and agreement factors. A computer program was written to do the calculation in projection using the same structure factor calculation as used in the systematic rotation described above (section 7.1). Intervals of $\frac{1}{8}$ hmax have been suggested as suitable (21) giving a grid of $1/48 \times 1/48$ of a cell edge in the present instance where hmax = kmax = 6. The calculation was carried out for angal in the orientation B derived from the sum function. The results are plotted in Fig. 21. The lowest minimum gave a structure very close to the correct structure. The three next lowest minima occurred at the symmetry related positions. Only one quarter of the unit cell needed to be covered. The plot shown was calculated using a corrected set of data but the calculation with the uncorrected data using 100.|E|values (or 200.|E| for the axial reflections) instead of reduced structure factors gave the best minimum in the same position.

7.4, SIGN RELATIONSHIPS

The signs of 21 hk0 reflections were derived from the triple product relationship (9):-

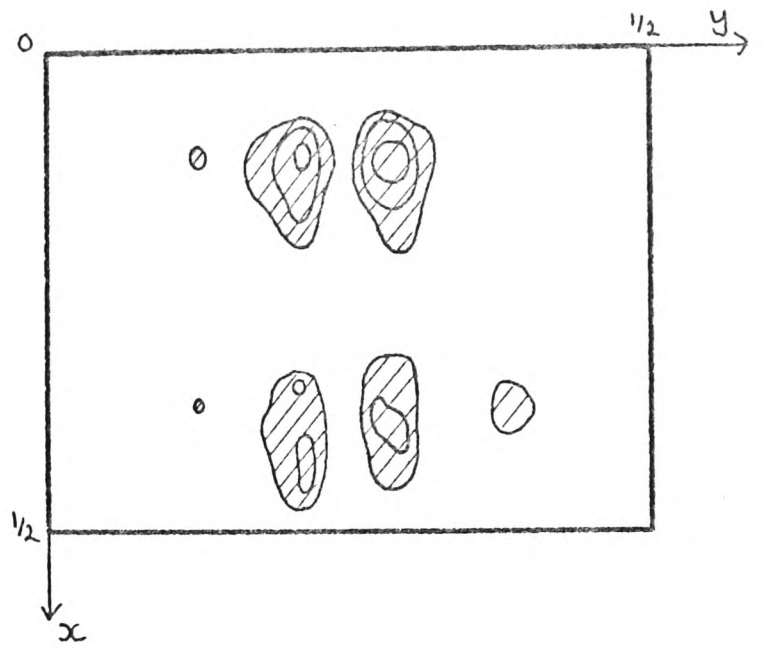
$s(\vec{h}) \cdot s(\vec{h'}) \cdot s(\vec{h+h'}) \approx 1$ for large values of $|E|_{\vec{h}}$, $|E|_{\vec{h'}}$, and $|E|_{\vec{h+h'}}$

$s(\vec{h})$ is the sign of the reflection $\vec{h}$ etc.

$\approx$ means probably equals.

The coincidence method of Grant, Howell and Rogers (37) was used. Reflections with $|E|>1.0$ were included in the

/calculation ...



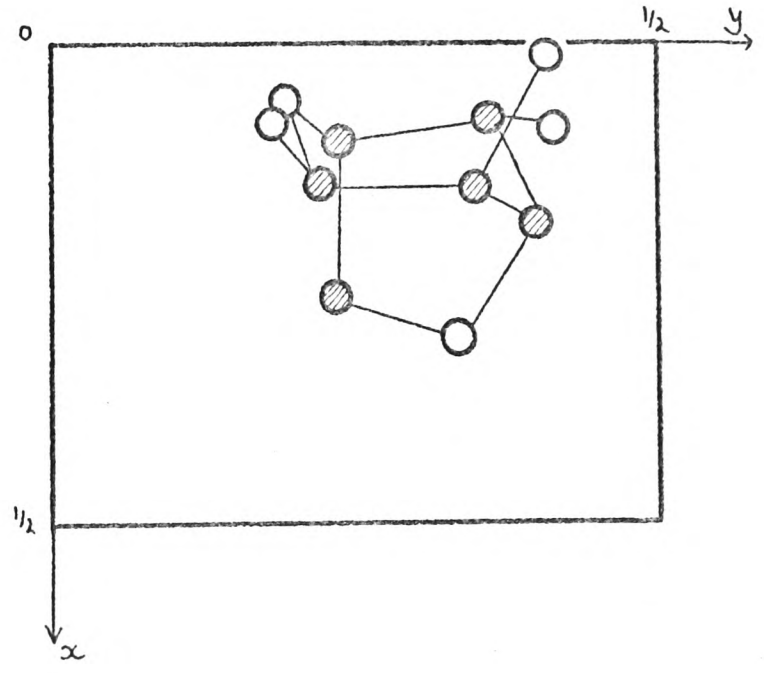


Fig. 21. (a) The best minima from the systematic translation of angal in orientation B derived from the rotation sum function and (b) the structure corresponding to the best minimum.
(Based on corrected intensity data.)

calculation. The probabilities of the various indications of signs were not calculated. An E-map was calculated and is shown in Fig. 22 together with two structures derived from it. The first of these gave the better fit to the E-map and gave good structure factor agreement at low sin θ . On attempting a least squares refinement however the atomic positions became meaningless though the R factor dropped to 0.27. The low sin θ reflections were assigned the highest weights in the hope that, if the structure were not quite correct, refinement would still be possible. The orientation of the six membered ring in the first structure was not consistent with the three dimensional Patterson function. In this respect the second structure was more promising and a possible hydrogen bond was noticed. This was between O(2) of one molecule and O(4) of a second molecule displaced by one cell translation parallel to the 'c' axis. The evidence in favour of this type of structure was fairly strong in spite of the poor structure factor agreement at low sin θ . An attempt was made to carry out a least squares refinement of this structure but it was not successful. The high weights assigned to the low sin θ reflections would probably account for this.

The errors in the high intensity reflections would probably have been detected more readily if it had not been for the fact that some of the trial structures gave good structure factor agreement for these high intensity reflections. At this stage in the structure determination no more work was done in projection

/as ...

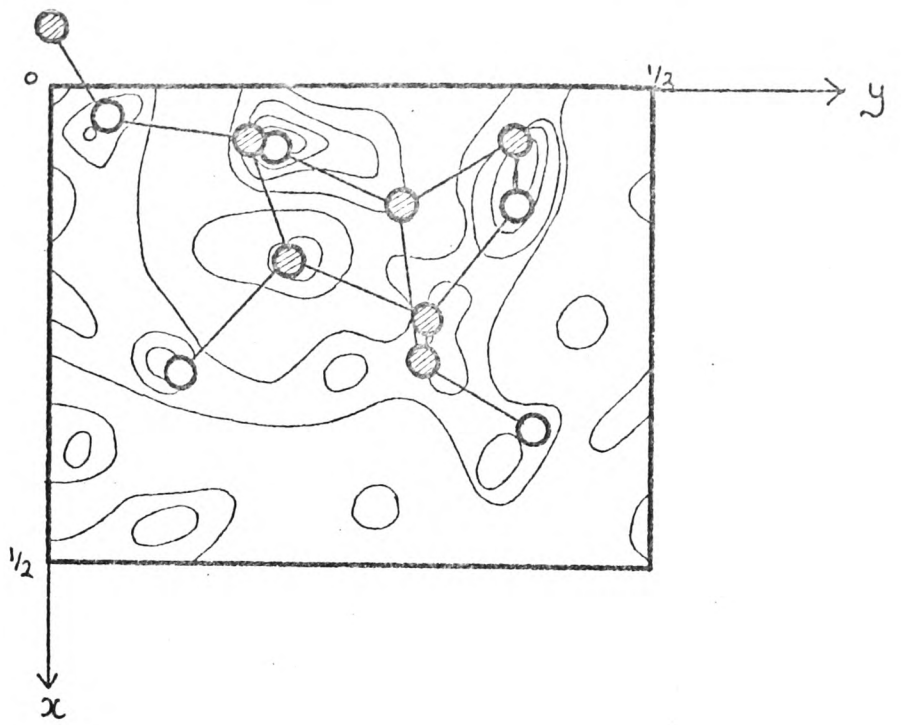


Fig. 22. E-map based on 21 two dimensional phases and two trial structures derived from it.

as the results were rather inconclusive and the derivation of a three dimensional set of phases, by direct methods, was being undertaken.

7.5, APPLICATION OF THE TANGENT FORMULA TO A FOUR ATOM PARTIAL TRIAL STRUCTURE

As a further test of the power of the tangent formula as a method of completing partial trial structures, a set of structure factors, based on a four atom partial trial structure, was calculated. The four atoms were placed at the peak positions numbers 5, 10, 6 and 11 of the E-map calculated from the first set of 60 phases derived by symbolic addition (Fig. 7). All four atoms were treated as being carbon atoms. 42 phases for which $|F|_{calc} > 0.4 \cdot |F|_{obs}$ and $|E| > 1.45$ were taken as the basic set. These were refined and the list of phases was expanded to 68 using the tangent formula. Only phases which had more than two indications and $|E|_{calc}$ values greater than 0.4 were accepted. Three cycles of tangent formula refinement of the basic set of phases and three cycles of refinement of the expanded set were carried out. The low number of phases obtained was due partly to the fact that once a phase had been rejected it was not reintroduced into the calculation. On another occasion provision would be made in the tangent formula program for reintroducing phases which had been rejected at earlier stages. An E-map based on the 68 phases was calculated and two projections of it are shown in Fig. 23. It showed peaks corresponding to

/eight ...

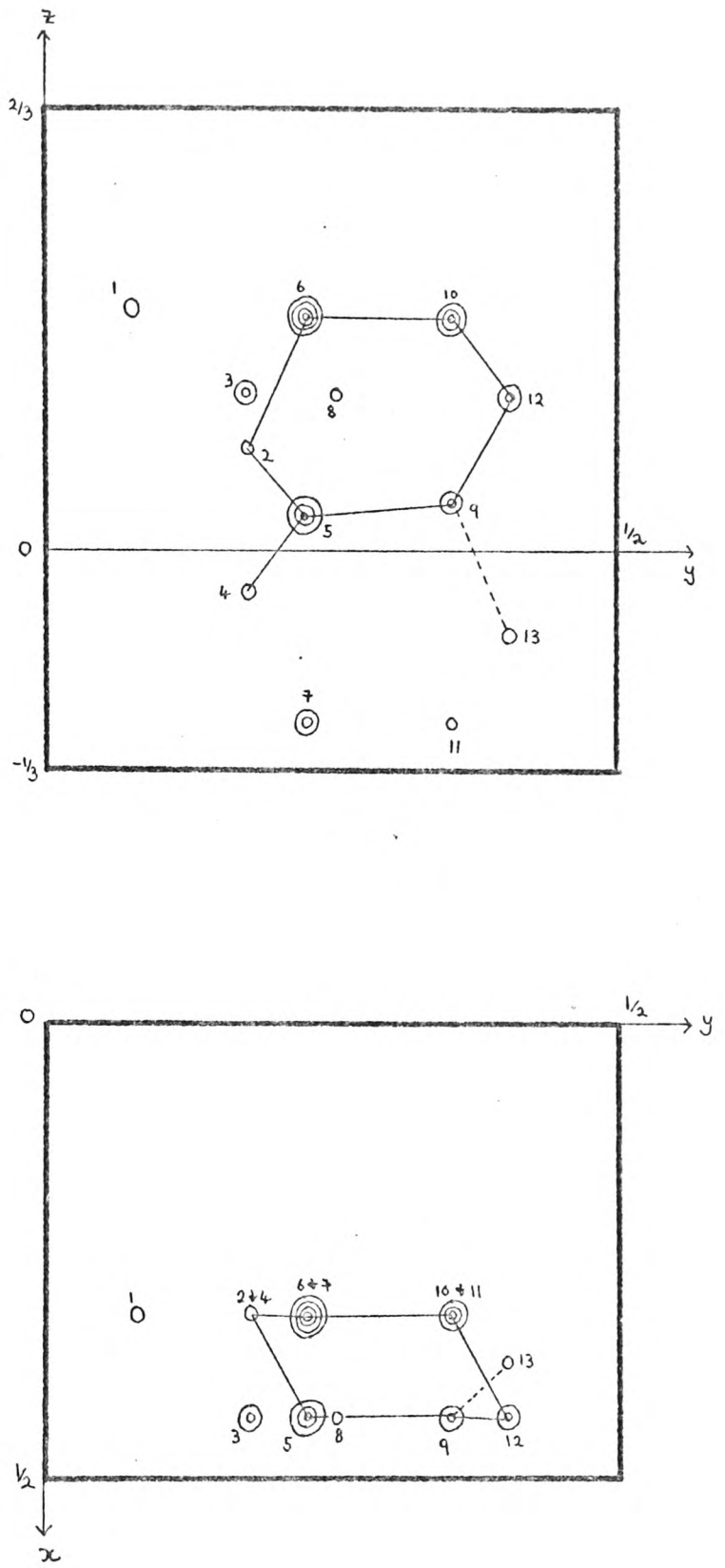


Fig. 23. Two projections (0kl, x=0 to x=$\frac{1}{2}$ and hk0) of E-map based on phases derived from a four atom partial trial structure.

eight atoms of angal, C(1), C(2), C(3), C(4), C(5), O(5) and O(2), though three of them were rather small and not very accurately placed. There were inaddition five spurious peaks of comparable size. Nevertheless the E-map was a distinct improvement on the four atom partial trial structure though in this case it was no more useful than the original 60 phase E-map. The 68 phases consisted of 21 two dimensional phases of which 18 were correct and 47 three dimensional phases with an average error of 46 degrees.

PART II

THE STRUCTURE OF IOTA CARRAGEENAN

SECTION I : INTRODUCTION

1.1, X-RAY DIFFRACTION OF POLYSACCHARIDES

The application of X-ray diffraction techniques to the structural investigation of polysaccharides requires a different approach from that used in single crystal structure determinations. The work must be carried out on oriented fibres of the polysaccharide which, even in the most favourable circumstances, give a very limited quantity of data. These fibres are most likely to be obtained when the polysaccharide is linear and has a regular repeating structure. It is almost essential that other evidence, including the knowledge of the primary structure of the polysaccharide derived by chemical means, be used in conjunction with the X-ray evidence. This is illustrated in a review by Marchessault and Sarko (38) on the "X-ray structure of polysaccharides". Most of the structures which have been investigated in some detail have been those of homopolymers such as cellulose, chitin and amylose.

The quantity which can most easily be derived from a fibre photograph is the fibre repeat distance which, combined with a knowledge of the primary structure of the polysaccharide, can give valuable information about the possible conformation of the chain. Further information may be derived from systematically absent reflections showing, for example, the presence of a two-fold screw axis in β-chitin (39) or from the general pattern of the intensity distribution due to a helical structure like amylose triacetate (40). Details of chain packing may also be deduced,/

deduced, though this is sometimes made more complicated by the presence of polymorphs.

At this stage models, which fit the general features of the diffraction pattern, must be made. They may be constructed either from physical models of various types or mathematically using an electronic computer. (See Section II on model-building). Consideration of steric factors, conformational energies and hydrogen bonding patterns can be used to predict possible conformations and packing arrangements. Conformational energies may be calculated approximately using various energy functions (68,69) or in some cases inferred from disaccharide crystal structure determinations. In the latter case it is possible that if crystal packing and hydrogen bonding are important factors in determining the conformation then these factors could be completely different for polysaccharide chains or in the presence of different solvents. Sundaralingham has done some work on predicting polysaccharide conformations from hydrogen bonding patterns (61). Experimentally derived evidence about hydrogen bonding patterns may sometimes be derived using polarised infra-red radiation techniques.

The most convincing argument for the correctness of a predicted structure would be close agreement between the observed and calculated intensity distributions. Refinement of the parameters of all the atoms in the repeating residue of a polysaccharide chain would be impossible with the quantity of data available from a fibre photograph. The refinement of a polysaccharide structure would best be done by carrying out a least squares refinement in which only parameters such as torsional angles/

angles around single bonds at glycosidic or side chain linkages and possibly bond angles at bridge oxygen atoms were varied, the geometry of the rigid groups being kept constant. This linked atom approach has been used in the refinement of the structures of the polypeptides α-poly-L-alanine (41,42) and β-poly-L-alanine (43) though, to date, there have been no applications of the method in the polysaccharide field. Unfortunately it is not always easy to calculate the intensity distribution especially when the fibre is semi-crystalline (44), when polymorphs co-exist or when there is unplaced solvent in the structure.

1.2, GEL FORMING POLYSACCHARIDES

Gel forming polysaccharides are found in many biological systems including animal connective tissues, bacterial capsules and the cell walls of growing plants (45). In the animal tissues gel forming polysaccharides are often associated with fibrous proteins such as collagen. Examples are chondroitin-4-sulphate which occurs in cartiledge and hyaluronic acid which is found in the vitreous humour of the eye and in synovial fluid surrounding joints.

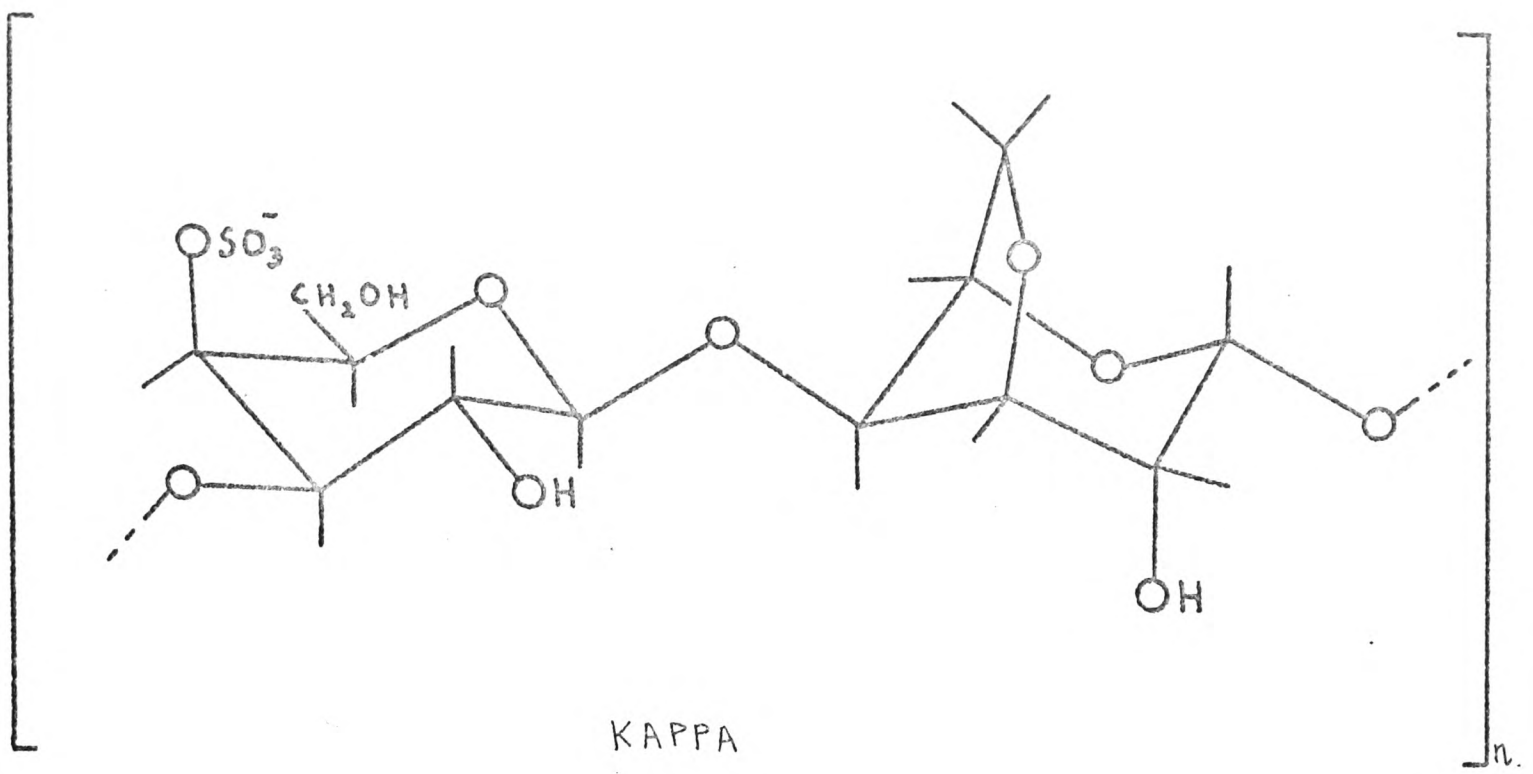
In grown plants the cell walls are based mainly on a matrix of cellulose fibres. In young plants however this would provide too rigid a cell wall and the cell walls are often based mainly on a matrix of a gel forming polysaccharide, usually a pectic substance, which can cross link to form a gel skeleton. This gel skeleton provides a less rigid cell wall which allows expansion as the young cells grow.

1.3, KAPPA AND IOTA CARRAGEENAN

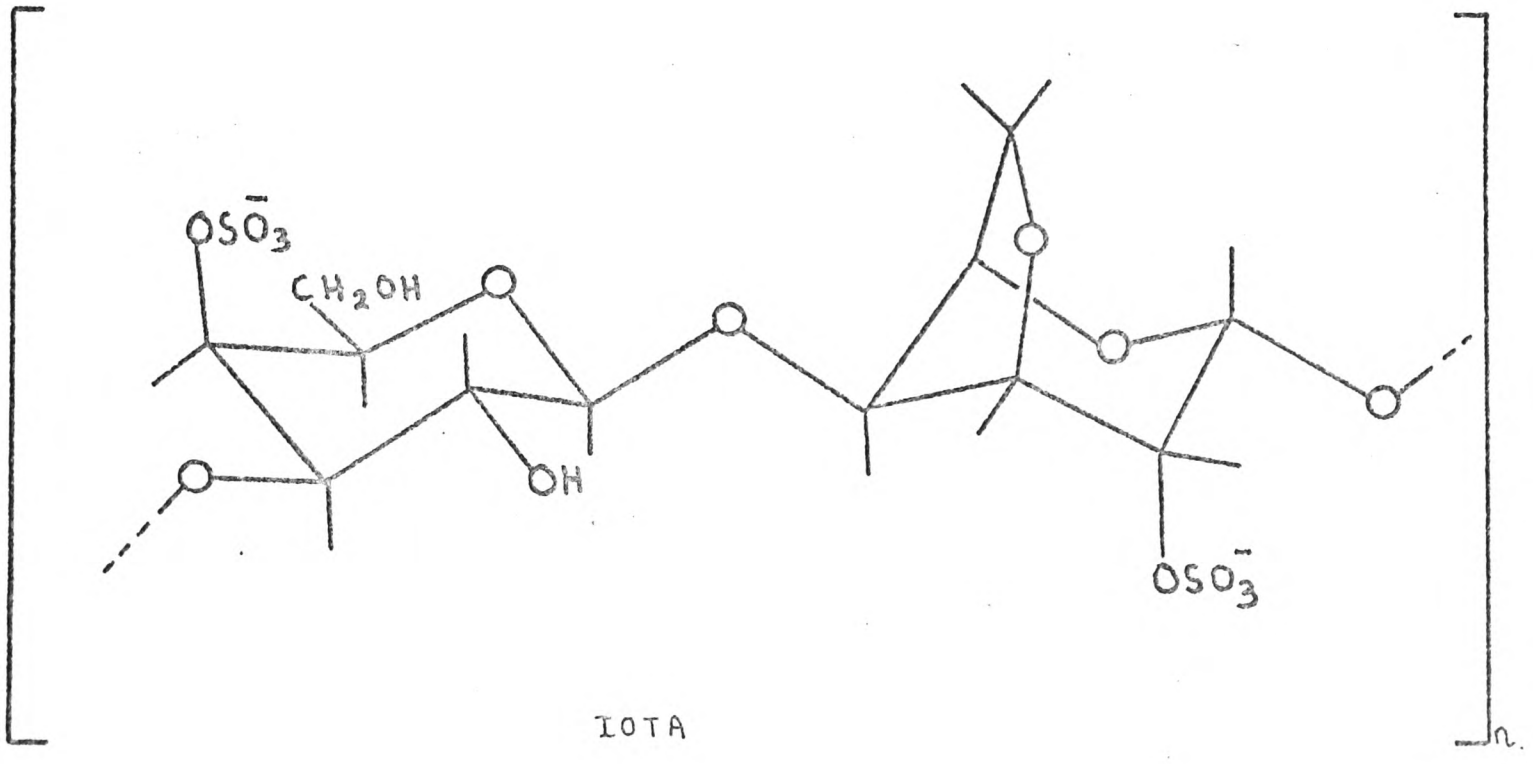
Seaweeds provide an important source of gel forming polysaccharides (46,47). Kappa and iota carrageenan are sulphated polysaccharides, based on galactose, which are extracted from certain red seaweeds. They interact selectively with cations to form strong gels. In their natural state they have masked repeating structures (48,49,50) but treatment with an alkaline solution of borohydride (51) converts them to regular repeating structures of the type $[-B-A-]_n$ where B is a 1,3 linked β-D-galactose-4- sulphate residue and A is a 1,4 linked 3,6-anhydro-α-D-galactose-2-sulphate residue in iota carrageenan and a 1,4 linked 3,6-anhydro-α-D-galactose residue in kappa carrageenan. These modified forms are henceforth referred to as iota and kappa (Fig.24). The sources and preparation of the polysaccharide samples used for the X-ray diffraction work are described by Anderson et al. (52).

1.4, KAPPA AND IOTA SALT FORMS

Different salt forms of kappa and iota were prepared by passing a dilute solution (approx. 0.5%) of the polysaccharide through an ion exchange column of a 100 - 200 fold excess of Amberlite IR 120 analytical grade resin in the appropriate salt form, followed by evaporation of the solution under reduced pressure at 35° C and freeze drying. Fibres were prepared by placing a drop of a warm aqueous solution of the polysaccharide (approx. 1-4%) between two glass beads in a cell of the type described by Fuller et al. (53). The glass beads were drawn apart/



KAPPA



IOTA

Fig. 24. The structures of kappa and iota.

apart to stretch the gel and the fibres were allowed to form by evaporation at 0 - 3° C for about 48 hours. Constant humidity was maintained by placing a saturated solution of sodium bromide within the cell.

Fibres of a series of monovalent cation salts of kappa were prepared and photographed by Dr. N.S. Anderson. A corresponding series for iota were prepared and photographed by Dr. J.W.B. Samuel and an unsuccessful attempt was made by him to obtain oriented fibres from some divalent cation salt forms of iota. Recently however Mrs. C. McNab was successful in preparing oriented fibres of the magnesium, calcium and strontium salts. The best fibres of iota were prepared from the potassium salt though good fibres were obtained with the lithium, sodium and the divalent cations. Rubidium was the largest cation which gave reasonably well oriented fibres. Kappa formed the best fibres with the larger cations. Oriented fibres of the lithium and sodium salts of kappa could not be prepared and that of the ammonium salt was rather poor.

1.5, FIBRE PHOTOGRAPHY

The photography by Dr. N.S. Anderson and Dr. J.W.B. Samuel of the monovalent cation salts was carried out using a Supper precession camera. Normally the fibre to film distance was 6.00 cms and the fibre was tilted out of the plane normal to the X-ray beam by 10 degrees. Cu Kα radiation from a Philips fine focus X-ray tube was used and this was collimated by a fine lead glass capillary. The fibre and the film remained stationary throughout/

throughout the exposure. The fibres of the divalent cation salts were photographed by Mrs. C. McNab using a Philips fibre camera with a fibre to film distance of 1.5 cms and also using a pin-hole camera constructed in the department. The best of the magnesium, calcium and strontium salt fibres were photographed by the author on the precession camera for the purpose of intensity measurement. In this case a collimator of the normal design was used as the use of a lead glass collimator did not appear to make much difference to the quality of the photographs though it did increase the required exposure times considerably.

1.6. KAPPA AND IOTA DIFFRACTION PATTERNS

The fibre repeat distances measured from the photographs were 22.4 ± 0.8 Å for kappa and 13.0 ± 0.3 Å for iota. The fibre repeat distance for iota was measured most accurately for the potassium salt though the salts with the other monovalent and divalent cations had very similar repeat distances. A determination of accurate spacings for all the iota salts is being carried out by Mrs. C. McNab using photographs calibrated with powder lines of known spacing. Small differences in spacings should be detected and any trends related to cation size should be revealed. The similarity of the diffraction patterns within each series of salts (Fig. 49) suggests that these series are iso-structural and that the structures of the divalent cation salts of iota are closely related to those of the monovalent cation salts. The diffraction photographs of kappa are much more diffuse than those of iota but the general intensity distribution on layers 0, 2, 4 and/

and 6 resembles that of layers 0, 1, 2 and 3 of iota. The first layer line in all the kappa salt photographs is very weak or absent. In both polysaccharides meridional reflections are only present when $l = 3n$ suggesting the presence of a three-fold screw axis.

1.7, UNIT CELLS

The positions of the first few orders of discrete reflections of the iota fibre photographs are consistent with a hexagonal packing arrangement with a = 22.6 Å. In all the salts the $10\bar{1}0$ reflection is absent and the first equational reflection is the $11\bar{2}0$ which corresponds to the first order of diffraction from the related hexagonal cell of side $a/\sqrt{3}$ = 13.0 Å (Fig.25). This is consistent with a hexagonal array of helices of diameter approximately 13 Å. A similar cell with a 20 Å is consistent with the kappa photographs. Density measurements (52) are consistent with three double helices (see below) with a water content of about 15% (54) passing through each unit cell.

1.8, DOUBLE HELIX MODELS

On the evidence outlined above a double helix model was proposed for both kappa and iota (52). In kappa each polysaccharide chain has three disaccharide residues in a turn of helix of pitch 24.6 Å and in iota three disaccharide residues in a turn of helix of pitch 26.0 Å. In iota the second chain is parallel to the first and displaced from it by half the helix pitch parallel to the helix axis, thus giving a crystallographic repeat/

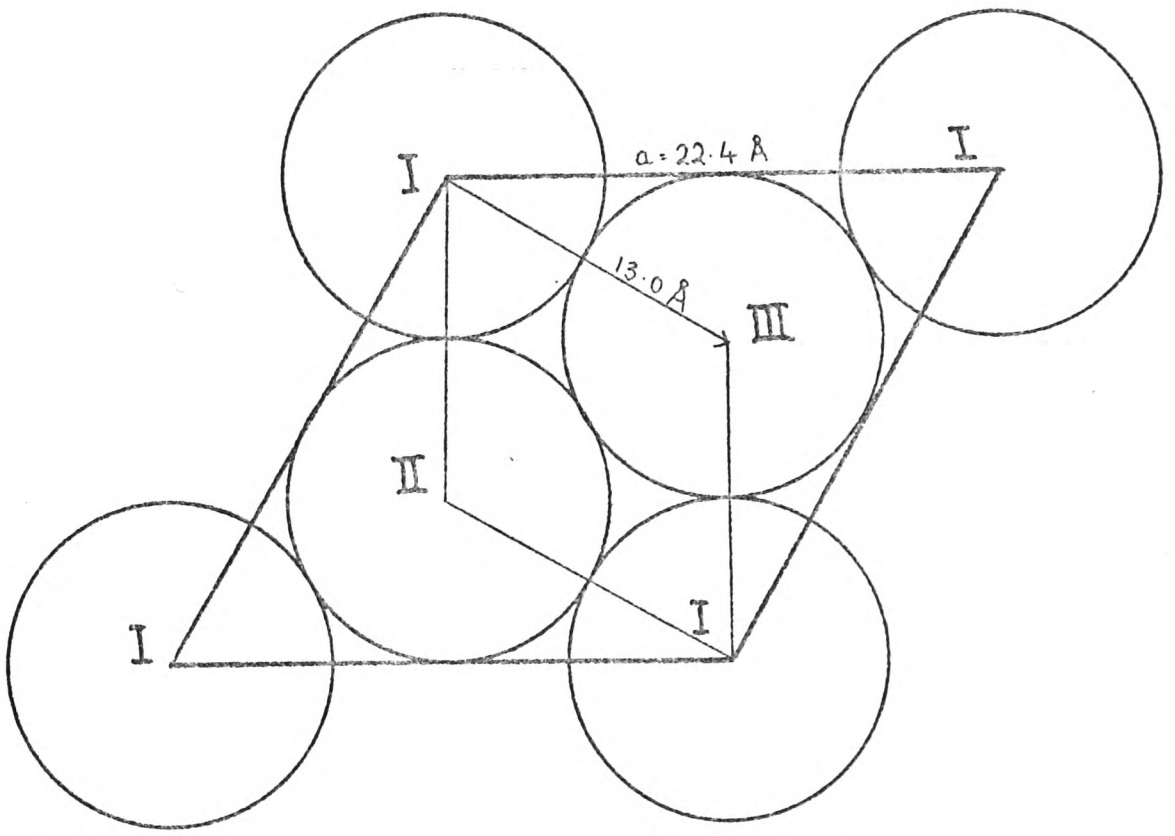


Fig. 25. The proposed hexagonal unit cell of iota showing the related hexagonal cell of side $a/\sqrt{3}$. Helices I, II and III are at different heights along the 'c' axis.

repeat distance of 13.0 Å. The relation between the chains in kappa is not yet known but the double helix model accounts well for the very weak first layer line. The reasonableness of the double helix model has been confirmed by model-building calculations and the 13 Å diameter for the double helix, deduced from the X-ray diffraction photographs, is consistent with the proposed model.

1.9, STRUCTURE AND FUNCTION

The proposed double helical model is the first example for a polysaccharide though a triple helix has been proposed for a β 1-3 linked xylan by Atkins et al. (55). The suggested mechanism for gel formation involves the formation of a network of polysaccharide chains which holds the water. The tie points between the polysaccharide chains are double helical regions. In later stages of gel formation these double helices probably aggregate in a similar manner to that observed in the stretched fibres. The fact that gelation is not prevented by the presence of different proportions of 4-sulphate on the galactose, or 2-sulphate on the 3,6-anhydrogalactose (56) whereas it is inhibited when some of the 3,6-anhydrogalactose units are replaced by galactose-6-sulphate units (57) is consistent with the proposed model. In the former case the chain conformation would be unaffected, whereas in the latter case a kink would be introduced into the chain wherever a 3,6-anhydrogalactose unit was replaced by a galactose-6-sulphate unit. Large changes in optical rotation as a 5.6% solution of the potassium salt of a segmented sample/

sample of iota carrageenan is heated and cooled over a range of 0 - 80° C can be attributed to double helix formation (58). The strength of the gels formed is dependent on the cation present and those which give the strongest gels also appear to give the best X-ray diffraction photographs. The effect of the cations, though probably important in the gel forming process, is not essential as neutral polysaccharides such as agarose, which is closely related to the carrageenans in structure, also form gels. At the present time the fibres obtained for kappa and iota are far superior in crystallinity and degree of orientation to those of agarose. Unfortunately it has not yet been possible to locate the cation positions.

SECTION II : THE FIBRE DATA

2.1, INTENSITY MEASUREMENTS

The intensity measurements of the fibre photographs were made using a Joyce - Loebl recording microdensitometer. Straight line traces were taken along the centre of each layer line and background traces were recorded on either side of the layer line. The two background traces were averaged and subtracted from the layer line trace to give the intensity distribution. No corrections were applied as the intensities measured from these traces were considered to be sufficiently accurate for calculating cylindrically averaged Patterson functions and for comparison with calculated cylindrically averaged intensity distributions for trial models. In the latter case the presence of any systematic error with theta could probably be noticed and partly compensated for by inclusion of an artificial temperature factor. For measurement of integrated intensities it would be important to apply at least a few of the largest corrections. The appropriate corrections for intensities measured from fibre photographs have been discussed by Franklin and Gosling (59).

Intensity traces were measured for the lithium, sodium, rubidium, magnesium and strontium salts of iota. Those of kappa and the ammonium and potassium salts of iota were measured by Dr. M.M. Harding who also calculated a chart for converting the x and z coordinates measured on the films for the tilted fibres to ξ and ζ coordinates. The intensity traces were redrawn on a rationalised ξ scale.

2.2,/

2.2, INDEXING OF STRONTIUM SALT PHOTOGRAPH

The positions of the discrete reflections on a photograph of a strontium salt fibre of iota, taken on a Philips fibre camera by Mrs. C. McNab, were measured using a travelling microscope. In this case the fibre axis was normal to the X-ray beam and the ξ and ζ coordinates were derived from the x and z coordinates using the expressions (60):-

$$\zeta = \frac{z}{\sqrt{(z^2 + x^2 + L^2)}}$$

$$\xi = \left\{ 2 - \zeta^2 - 2\sqrt{(1-\zeta^2)} \cdot \frac{L}{\sqrt{(L^2+x^2)}} \right\}^{1/2}$$

L = fibre to film distance.

The fibre to film distance was not known very accurately and the ξ and ζ coordinates were scaled such that the side of the hexagonal cell a was 22.4 Å. For the purposes of indexing this was unimportant as only the relative coordinates need be known. The results are shown in Table 7. The positions of the reflections on the equator are consistent with a hexagonal arrangement where the first observed reflection is the $11\bar{2}0$ and where $-h + k + l = 3n$ is a condition of reflection. On the upper layers the positions of the reflections are consistent with the hexagonal arrangement but $-h+k+l = 3n$ is no longer a condition of reflection. The results are very similar to those found for the monovalent cation salts. For these the unit cell shown in Fig. 25 was proposed (52) where helices I, II and III were at different heights in the unit cell. The related hexagonal cell of side $a/\sqrt{3}$ is also shown. It is difficult to imagine how any such arrangement could be formed by three-fold helices unless the heights of I, II/

l=0

-h+k+l=3n
a condition

NO.	ξobs.	ξcalc.	indices
1	0.136	0.138	$11\bar{2}0$
2	0.237	0.239	$03\bar{3}0$
3	0.278	0.276	$22\bar{4}0$
4	0.369	0.364	$14\bar{5}0$
5	0.515	0.496	$25\bar{7}0$
		0.550	$44\bar{8}0$
6	0.633	0.631	$36\bar{9}0$

l=1

		-h+k+l=3n a condition		no conditions	
NO.	ξobs.	ξcalc.	indices	ξcalc.	indices
1	0.076	0.079	$10\bar{1}1$	0.079	$10\bar{1}1$
2	0.135	0.158	$02\bar{2}1$	0.138	$11\bar{2}1$
3	0.237	0.209	$21\bar{3}1$	0.239	0331
4	0.276	0.287	$13\bar{4}1$	0.274	$22\bar{4}1$
5	0.369	0.397	$05\bar{5}1$	0.364	$14\bar{5}1$
6	0.423	0.422	$24\bar{6}1$	0.422	$24\bar{6}1$

l=2

		-h+k+l=3n a condition		no conditions	
NO.	ξobs.	ξcalc.	indices	ξcalc.	indices
1	0.071	0.079	$10\bar{1}\bar{2}$	0.079	$10\bar{1}\bar{2}$
2	0.138	0.158	$02\bar{2}\bar{2}$	0.138	$11\bar{2}2$
3	0.202	0.209	$21\bar{3}\bar{2}$	0.209	$21\bar{3}\bar{2}$
4	0.239	?		0.239	$03\bar{3}2$
5	0.281	0.287	$13\bar{4}\bar{2}$	0.287	$13\bar{4}\bar{2}$
		0.274	$12\bar{3}2$	0.274	$12\bar{3}2$
6	0.341	0.346	$32\bar{5}\bar{2}$	0.346	$32\bar{5}\bar{2}$
7	0.400	0.397	$05\bar{5}\bar{2}$	0.397	$05\bar{5}\bar{2}$

Table 7, The indexing of the reflections of the strontium iota fibre photograph. (Based on a hexagonal cell with a = 22.4Å.)

II and III were 0, 1/3 and 2/3 respectively. If this were so then the space group would be R3 and $-h + k + l = 3n$ would be a condition of reflection throughout. It would seem reasonable therefore to suppose that the packing is based on this space group+ but that shift distortions (i.e. displacements of the molecules parallel to the helix axis) about the mean positions of 0, 1/3 and 2/3 are present thus allowing 'forbidden' reflections to appear on the upper layers though not on the zero layer. That shift disorder is present is indicated by streaking along the layer lines of the upper layers. In the strontium salt there is fairly strong streaking along the zero layer as well indicating that lateral distortion is present. If the layer line at which the intensity distribution became continuous could be observed then the amount of shift distortion could be calculated (60). This layer line is however beyond the fifth which is the maximum normally present in the iota photographs. It should be noted that if the shift was continuous then discrete reflections would only be observed on the equator and the side of the unit cell would be that of the projected cell, 13Å. Possible arrangements involving anti-parallel packing of the iota double helices are being considered by Dr. M.M. Harding. Packing arrangements, with displacements/

+Note: A packing arrangement that would account for the observed absences would be a mixture of space groups R3 and $P3_2$. The side of the trigonal cell would be 13 Å. This would account for all the reflections measured for the strontium salt though not for the off meridional reflection on the third layer line of the monovalent cation salts. Indexing on the basis of the hexagonal cell would give $-h + k + l = 3n$ as a condition of reflection on layer lines where $l = 3n$ but would give apparently general hexagonal reflections on the other layers.

displacements of the anti-parallel double helices relative to each other, could account for the streaking along the layer lines and, in view of the conditions of formation of the polysaccharide fibres, such packing could reasonably be expected to occur.

2.3, COMPARISON OF MONOVALENT AND DIVALENT SALT DIFFRACTION PATTERNS

In the potassium salt fibre photograph of iota the position of the second reflection along the equator is not consistent with the indexing scheme described above or even for one allowing all hexagonal reflections for a cell of side 22.4 Å. The other equatorial reflections appear to follow the same pattern as that of the strontium salt though some of them are rather weak and their positions are not easily measured. The equivalent reflection is also present in the lithium and magnesium salt photographs and could possibly be due to the particularly high value of the molecular transform at that point (see p.104). On inspection of the microdensitometer traces of the strontium salt photograph a slight maximum between the first and second equatorial reflections was observed though this was rather broader than the discrete reflections that were measured using the travelling microscope. Again this suggests that the maximum at this point may be due to the large magnitude of the Fourier transform.

One interesting difference between the monovalent and divalent cation salt photographs of iota occurs on the third layer line. In the monovalent cation salt fibres there is a strong/

strong reflection just off the meridian as well as the strong meridional reflection. This reflection is difficult to explain in terms of the proposed structure as the calculated value of the transform at this point is rather low compared with the meridional maximum (see Fig.44). It is also in the position of a 'forbidden' reflection by the $-h + k + l = 3n$ condition. A photograph, taken by Dr. J.W.B. Samuel, supposedly of a sodium salt fibre was found to have these off meridional reflections strong and the meridional reflection weak. In this photograph the positions of the low order reflections were different from the other photographs indicating the presence of another crystal form. Most of the sodium salt photographs followed the normal pattern for the monovalent cation salts but it would seem possible that more than one packing arrangement might be present in the monovalent cation salt fibres. In contrast these off meridional reflections were absent or at least very weak compared with the meridional reflections in the case of the divalent cation salts.

Photographs of the lithium, potassium, magnesium, strontium and the two types of sodium salts are shown in Fig.49.

SECTION III : MODEL-BUILDING

3.1, MODEL-BUILDING

In attempting to derive sterically feasible conformations for a polysaccharide chain, mathematical model-building procedures for use on an electronic computer have advantages over working with physical models. The mathematical model provides an easily adjustable model and using a high speed computer it is possible to investigate systematically a wide range of conformations in a reasonable amount of time. The accuracy of a mathematically built model is much greater than that of a physical model, being in general only limited by the accuracy of the input data. Mathematical model-building methods may be used to predict possible conformations of a polysaccharide chain or for providing a variable model which can be fitted to some experimentally derived data. A combination of these two approaches may also be used.

3.2, ASSUMPTIONS

Model-building methods for polysaccharides, whether they use molecular models or the computer, are based on the assumption that the geometry of the individual monosaccharide units can be regarded as fixed and that it can either be derived from a solid state crystal structure determination of the monosaccharide in question or be derived by assuming standard values for bond lengths and bond angles. The bridge oxygen angles may also be regarded as fixed, though there is less information available

about/

about the best values to use. Angles in the range 115 - 120° would seem reasonable as the bridge oxygen angles in the disaccharides, whose crystal structures have been determined, fall within this range. Typical bond lengths and bond angles for estimating a set of monosaccharide coordinates are discussed in Part I, section 6.1 of this thesis. The only variable parameters affecting the conformation of the main backbone of the polysaccharide chain are then the angles of rotation (torsional angles) around the carbon to bridge-oxygen bonds.

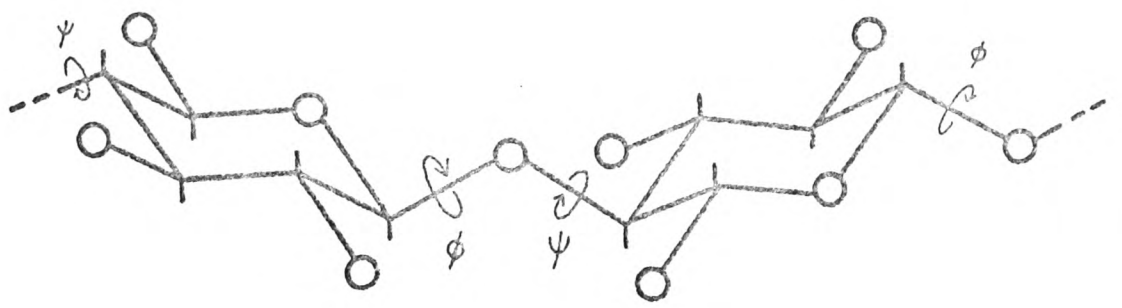
3.3, REPEATING STRUCTURES

The simplest repeating structure is one of the type $[-A-]_n$, for example cellulose (Fig.26). If the conformation at each linkage along the chain is the same then the conformation of the polysaccharide chain can be defined in terms of two variable parameters, the torsional angles ϕ and ψ as shown. In the case of the polysaccharides kappa and iota the repeating structure is of the type $[-A-B-]_n$. Here there are two different linkages present and thus four variable parameters ϕ_1 and ψ_1 at the B-A linkage and ϕ_2 and ψ_2 at the A-B linkage (Fig.26). Again the assumption is made that the conformations at the two linkages are repeated along the length of the chain.

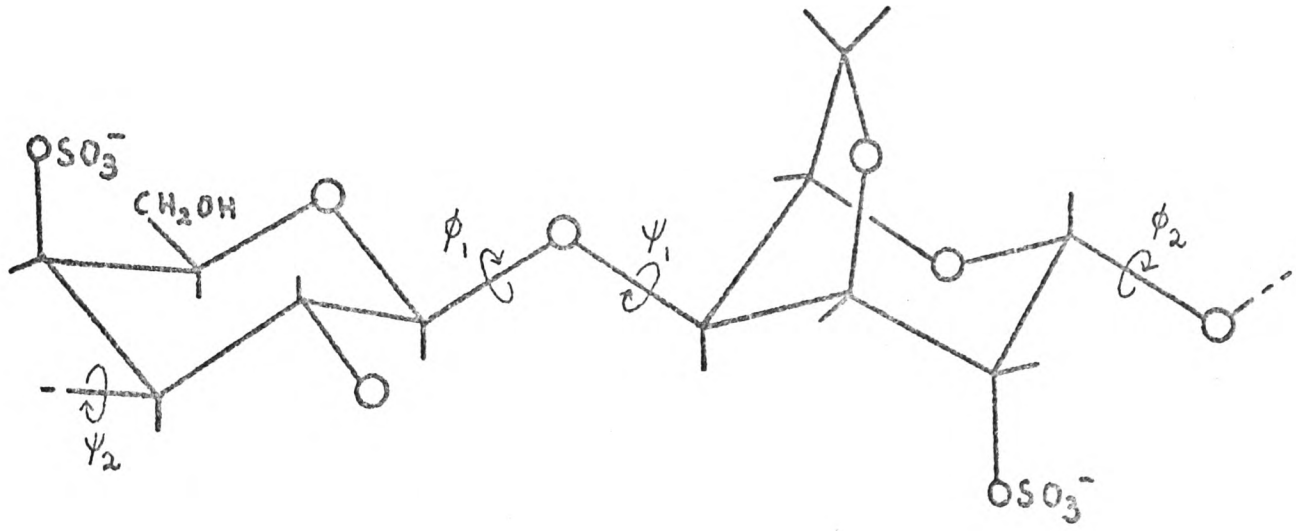
3.4, HELIX PARAMETERS

Such repeating structures will in general form helices. Two important helical parameters are $\underline{h}^+$ the pitch of the helix and/

+Note. This is not the same as the helix parameter h used by Rees (62) and by Ramachandran et al.(32) which refers to the length of the repeating residue projected onto the helix axis.



a. cellulose.



b. iota.

Fig. 26. Polysaccharide repeating structures of the types $(-A-)_n$, e.g. cellulose and $(-B-A-)_n$, e.g. iota showing the variable torsional angles.

and n the screw symmetry or the number of residues per turn of helix. In the case of repeating structures of the type $[-A-B-]_n$ a residue is a disaccharide unit. From the values of the four torsional angles it is possible to calculate the corresponding helix parameters.

3.5, STERIC CONSIDERATIONS

In considering possible conformations for a polysaccharide chain, those conformations which are not sterically feasible are rejected. Possible conformations at each of the linkages may be considered separately. The torsional angles ϕ and ψ at the linkage are varied stepwise over 360 degrees and for each value of ϕ taken with each value of ψ the distances between the atoms of adjacant monosaccharide units are calculated. If the distance between any pair of atoms is less than the 'allowed' van der Waals contact distance for that particular pair then the conformation is disallowed. If there are no such short contacts then the conformation is considered further. The results of such a calculation may be plotted out in the form of steric maps, examples of which are shown in Figs. 33 and 34.

3.6, MODELBUILD I AND MODELBUILD II

Two mathematical model-building approaches have been applied to the iota structure. In the first of these, modelbuild I, the two experimentally determined helical parameters h and n were used to fix two of the four variable parameters of the polysaccharide chain. The allowed conformations at the B-A linkage/

linkage were calculated and for each allowed conformation the values of ϕ_2 and ψ_2 were calculated from ϕ_1 , ψ_1 , h and n. In general, for each pair of values of ϕ_1 and ψ_1 there were four possible solutions for ϕ_2 and ψ_2 , two corresponding to right handed helices and two corresponding to left handed helices. The values of ϕ_2 and ψ_2 were inspected to see whether they gave allowed conformations at the A-B linkage and if they did they gave sterically feasible solutions for a single polysaccharide chain having the required helix parameters h and n.

Modelbuild I was used to predict possible conformations of iota assuming the correctness of the double helix model with a pitch of 26.0 Å. Any solutions for the conformation of the single polysaccharide chain were further investigated to see whether it was sterically possible to fit in a second chain in the symmetry related position.

This approach has the advantage that a large amount of computing time is saved as only two parameters have to be systematically varied instead of four and only those conformations which fit the experimentally derived helical parameters are considered. This approach cannot however be used for polysaccharides which do not form sufficiently well oriented fibres for the helical parameters to be determined experimentally. To approach the problem from a quite general point of view a second method, modelbuild II, was developed by Dr. D.A. Rees (52). Only stereochemical restrictions were considered at first. Each allowed conformation at the A-B linkage was taken with each allowed conformation at the B-A linkage and for each combination various properties of the chain were calculated, for example,/

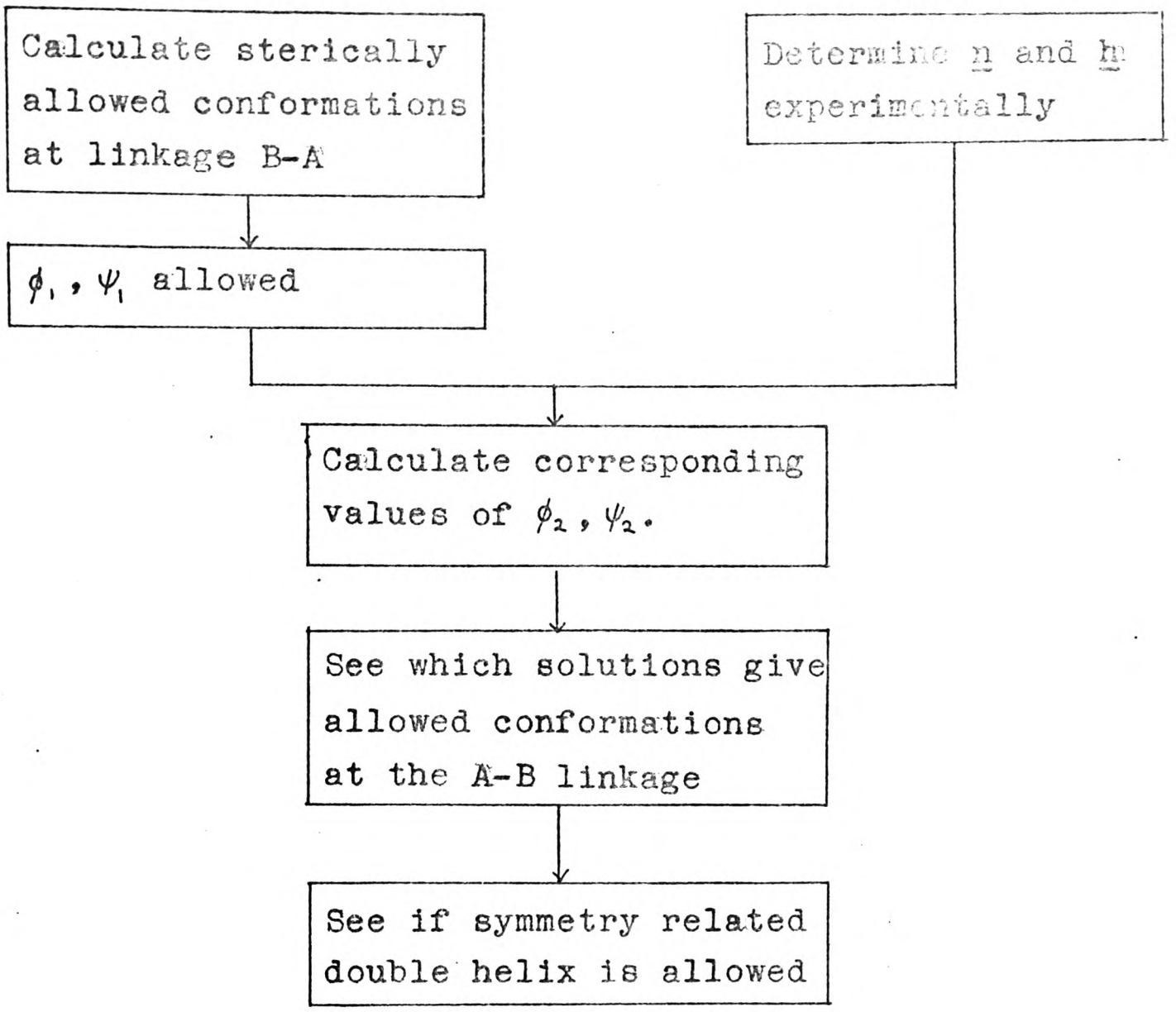
example, the screw symmetry and the length of the disaccharide residue projected onto the helix axis. The possibility of forming double helices was also investigated. Though in iota the X-ray diffraction evidence indicates that the second chain of the double helix lies parallel to the first and is displaced from it by half a repeat distance, other double helices may have the chains related in less simple ways. The chains could, for example, be anti-parallel. Modelbuild II has been applied to iota carrageenan and also to a number of other polysaccharides (62).

The two model-building procedures are summarised in Fig.27.

3.7, MODELBUILD I

Figs. 28 and 29 illustrate in more detail the method of calculation for modelbuild I. A pair of values ϕ_1 and Ψ_1, which gives an allowed conformation at the B-A linkage, is selected. The length of the disaccharide from bridge oxygen to bridge oxygen is calculated for this conformation. For iota this was the distance between O(3) of the galactose (B) and O(1) of the 3,6 anhydrogalactose (A). The positions of these O(1/3) atoms are fixed at points P_i on a helix of the required h and n (Fig.29). It is now possible to rotate the disaccharide residue I around P_1 P_2 without changing its conformation or the helix parameters. As residue I is rotated around P_1 P_2 and residue II is rotated around P_2 P_3 such that its position is related to the first by the symmetry of the helix, the bridge oxygen angle at P_2 varies. There will be two values of this angle of rotation Φ which give any particular value of the bridge oxygen angle/

MODELBUILD I



MODELBUILD II

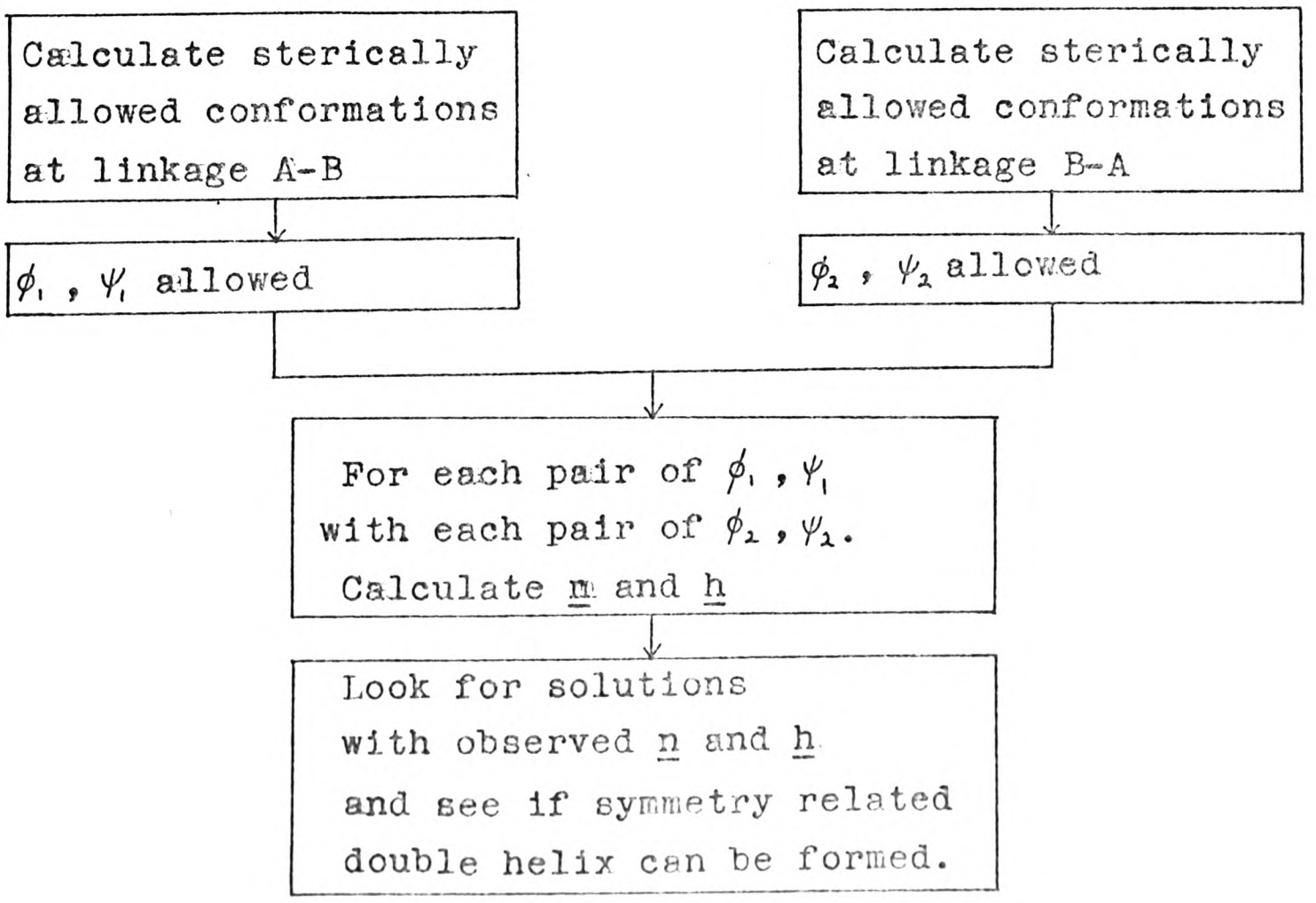


Fig. 27, Outline of modelbuild I and modelbuild II as applied to the structure of iota.

MODELBUILD I

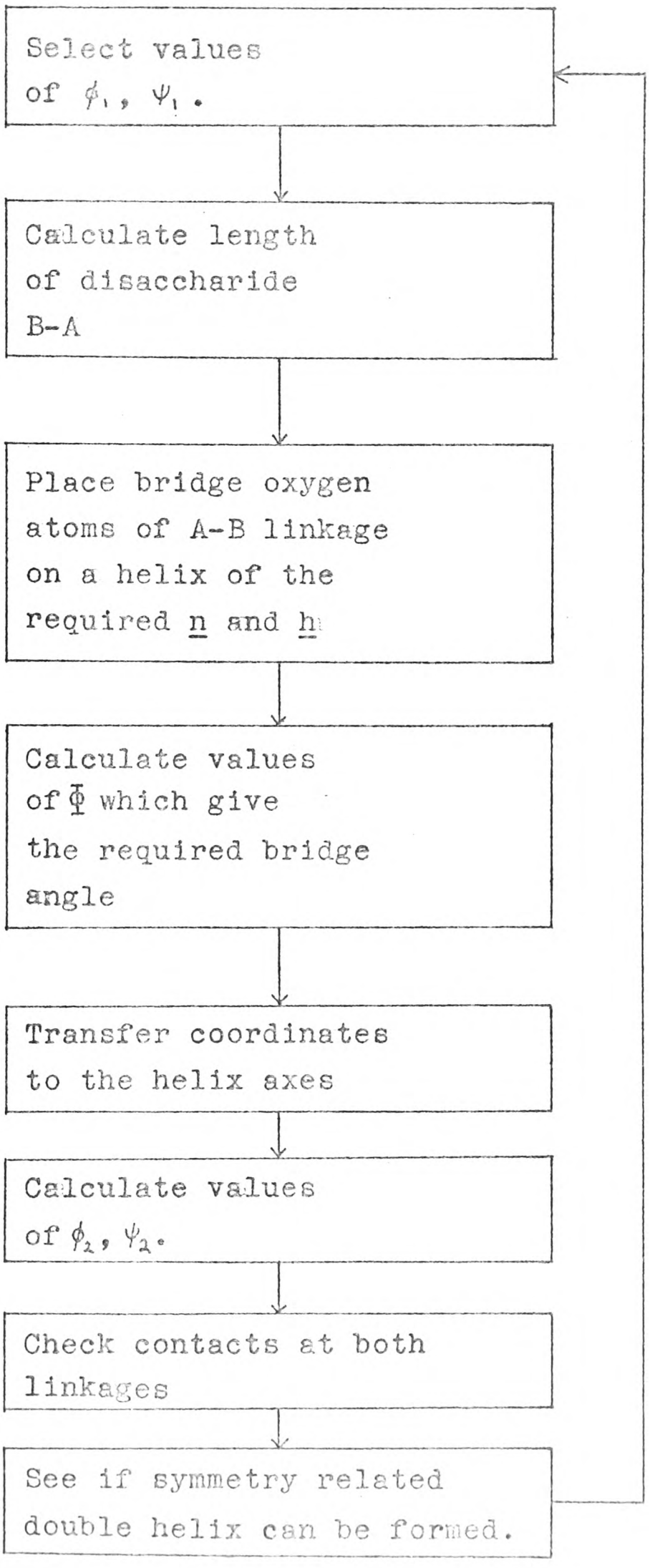


Fig. 28, Outline of modelbuild I calculation as applied to the structure of iota.

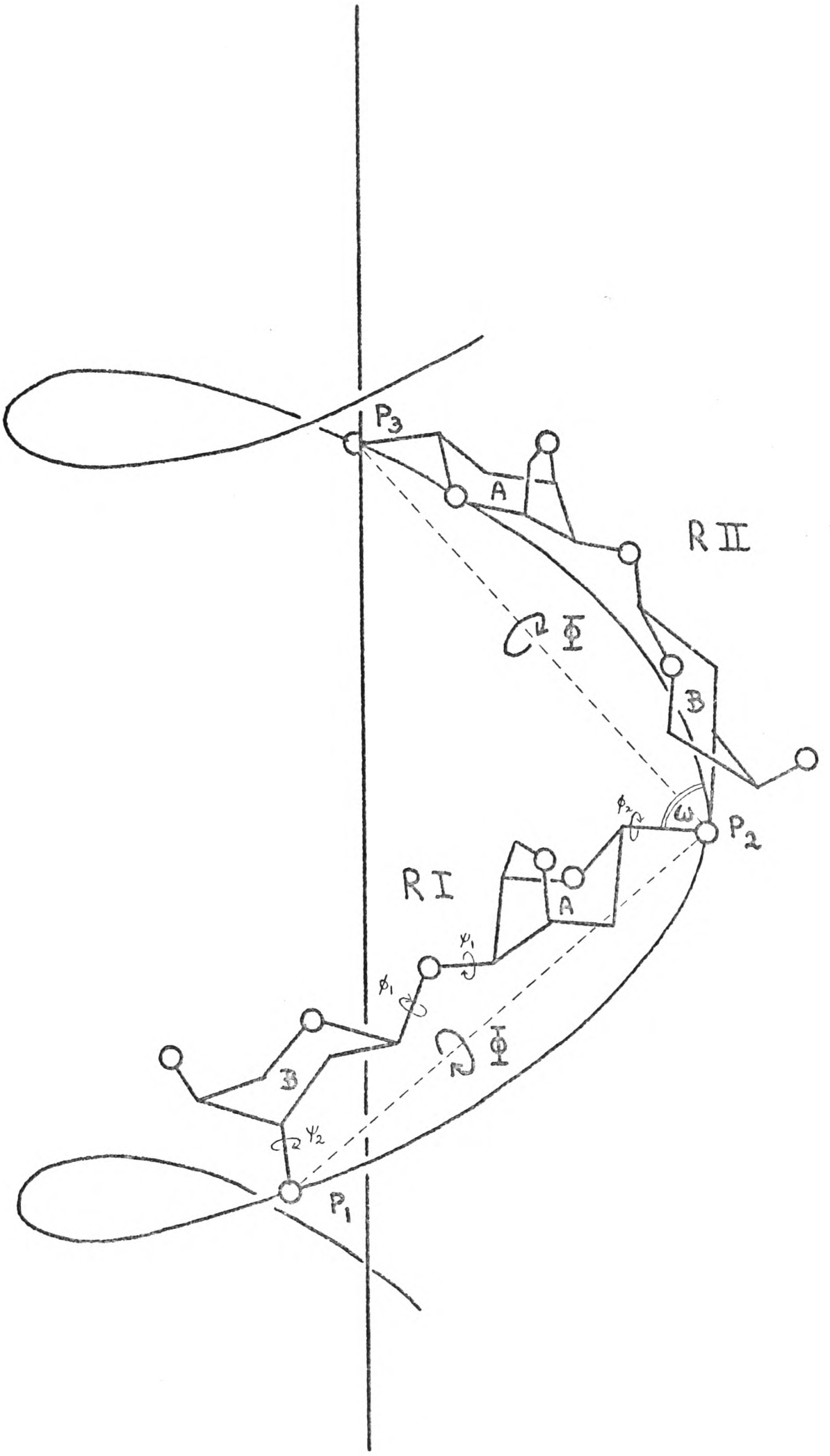


Fig. 29. The geometry of modelbuild I showing the O(1/3) atoms of the A-B linkage fixed on the required helix.

angle ω. Alternatively, if a value is assigned to ω there will in general be two possible solutions for Φ which give the required value of ω, though in some cases there will be no values of Φ which give the required value of ω. In the latter case the required helix cannot be constructed for the pair of values of ϕ_1 and ψ_1 which were selected. If there are solutions then the coordinates of the disaccharide residue are calculated with respect to the helix axes and the torsional angles ϕ_2 and ψ_2 are calculated. Then, as described above, these values are inspected to see whether they give allowed conformations at the A-B linkage and the possibility of forming the symmetry related double helix is investigated. The process is repeated for the next pair of values of ϕ_1 and ψ_1 until all the allowed conformations have been covered. Right handed and left handed helices are treated separately and the complete calculation is carried out for each case.

A similar approach has been used by Jones (63) for cellulose, by Settineri and Marchessault (64) for xylan and by Sarko and Marchessault (65) for amylose triacetate, though in these cases the repeating structures were of the simplest kind $[-A-]_n$. Their approach also differed in that the angle of rotation Φ was treated as a variable parameter as far as the calculations were concerned. Instead of solving an equation to find Φ, as in modelbuild I, Φ was systematically varied and ω was calculated as a function of Φ. This function was plotted out and the values of Φ which gave chemically sensible values for ω were found by inspection. This has the advantage that a range of values of ω is allowed for but/

but in the case of repeating structures of the type $[-A-B-]_n$ the introduction of Φ as a third variable parameter would be undesirable. If it were considered necessary, the modelbuild I calculation could be repeated for a series of values of ω. When large numbers of conformations are being considered it would be very time consuming to vary a third parameter and to inspect the results for possible solutions.

3.8, THE MATHEMATICS OF MODELBUILD I

The coordinates of monosaccharide A with respect to axes AX, AY, AZ (Fig.30a) and of monosaccharide B with respect to axes BX, BY, BZ were input. The zero position for the torsional angle ϕ_1 around the C(1) - O(1) bond was defined by choosing an atom of B to lie in the BX-BZ plane with its BZ coordinate positive when $\phi_1 = 0$. For iota C(4) of galactose was chosen to define the zero position of ϕ_1. Similarly C(1) of the 3,6-anhydrogalactose was chosen to lie in the AX-AZ plane with its coordinate positive when $\psi_1 = 0$.

A positive rotation of ϕ_1 or ψ_1 was taken as being a clockwise rotation of the monosaccharide residue around the bridge-oxygen to carbon bond looking along the direction from the bridge oxygen towards the appropriate monosaccharide residue.

For each pair of values of ϕ_1 and ψ_1 the coordinates of residue A were transformed to axes BX, BY, BZ by the operation of three matrices M1, M2 and M3 where:-

(i) M1 was a rotation around the AX axis by $-\psi_1$

(ii) M2 was a rotation around the new 'y' axis by $-\alpha'$ (α' was the/

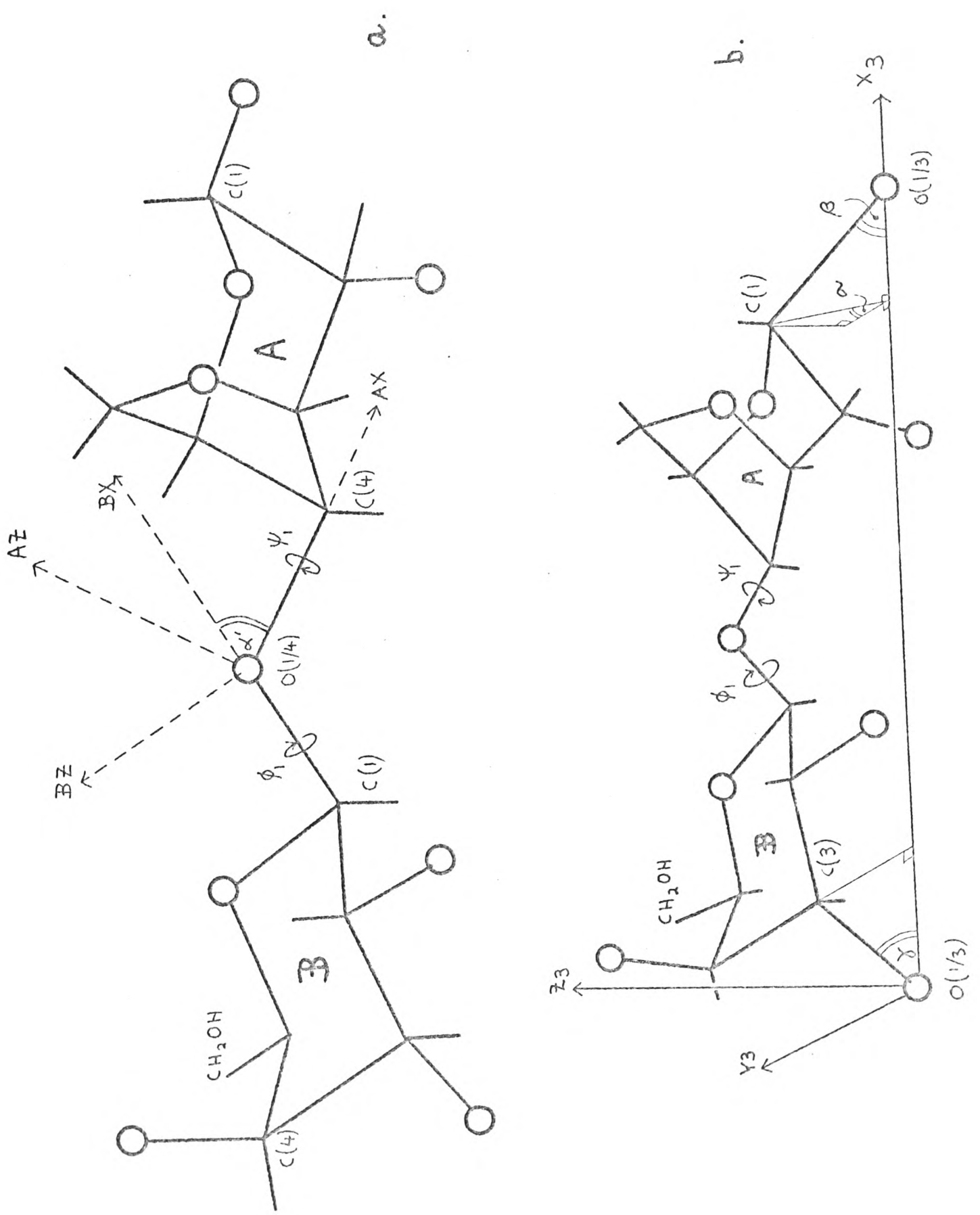


Fig. 30. The sets of coordinate axes on the disaccharide residue B-A used in modelbuild I.

the supplement of the bridge oxygen angle)

(iii) M3 was a rotation around the new 'x' axis by $-\phi$.

The three matrices were:-

$$M1 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & -\sin\psi \\ 0 & \sin\psi & \cos\psi \end{vmatrix}$$

$$M2 = \begin{vmatrix} \cos\alpha' & 0 & \sin\alpha' \\ 0 & 1 & 0 \\ -\sin\alpha' & 0 & \cos\alpha' \end{vmatrix}$$

$$M3 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{vmatrix}$$

The coordinates of the disaccharide residue were transformed to a set of axes X3, Y3, Z3 as shown in Fig.30b.

The line joining the bridge oxygen atoms of the A-B linkage defined the X3 axis. The Y3 axis was chosen such that C(3) of the galactose lay in the X3-Y3 plane with its Y3 coordinate positive. The length of $\underline{l}$ of the disaccharide residue was calculated. If this was less then $\underline{h}/\underline{n}$ then obviously the required helix could not be formed for that conformation of the disaccharide.

The angles γ , β and α were calculated where γ was the angle between the O(3) - C(3) bond of galactose and the X3 axis, β was the supplement of the angle between the O(1) - C(1) bond of the 3,6-anhydrogalactose and the X3 axis and α was the positive angle of rotation around the X3 axis of C(1) of the 3,6-anhydrogalactose, the zero position being taken in the X3-Y3 plane/

plane with the Y3 coordinate positive.

P_o, P_1, P_2 etc. (Fig.31) were the positions of the A-B linkage bridge oxygen atoms placed on a helix of the required $\underline{h}$ and $\underline{n}$. The angle Θ between $P_o - P_1$ and $P_1 - P_2$ was calculated from the expression:-

$$\cos\Theta = \frac{1}{2\underline{\ell}^2}\left\{4\frac{\underline{h}^2}{\underline{n}^2} + 2\left(\underline{\ell}^2 - \frac{\underline{h}^2}{\underline{n}^2}\right)\right\}\left\{1 + \cos\left(\frac{2.\pi}{\underline{n}}\right)\right\} - 1$$

The zero position of $\underline{\Phi}$, the angle of rotation of residue II around P_2 P_3, was taken as being the position where C(3) of the galactose of residue II lay in the P_1 P_2 P_3 plane. The symmetry related residue I has C(3) of galactose rotated by $\underline{\Phi}$ out of the P_o P_1 P_2 plane or $\underline{\Phi}-\delta$ out of the P_1 P_2 P_3 plane when δ is the angle between the planes P_o P_1 P_2 and P_1 P_2 P_3.

The angle δ was calculated from the expression:-

$$\cos\delta = 1 + \frac{\underline{\ell}^2(1+2\cos\Theta)^2 - (P_o P_3)^2}{2\underline{\ell}^2\sin^2\Theta}$$

where $$(P_o P_3)^2 = 3\frac{\underline{h}^2}{\underline{n}^2} + \left\{\underline{\ell}^2 - \frac{\underline{h}^2}{\underline{n}^2}\right\}\left\{\frac{\sin^2(3\pi/\underline{n})}{\sin^2(2\pi/\underline{n})}\right\}$$

The detailed geometry of the bridge oxygen atom at P_2 is shown in Fig.32. Points Q_1 and Q_2 were taken along the O(1/3) - C(1) bond of the 3,6-anhydrogalactose and along the O(1/3) - C(3) bond of the galactose at unit distance. The required value of the bridge oxygen angle was ω.

Then $$\frac{(Q_1 Q_2)^2}{2} = 1 - \cos\omega$$

A/

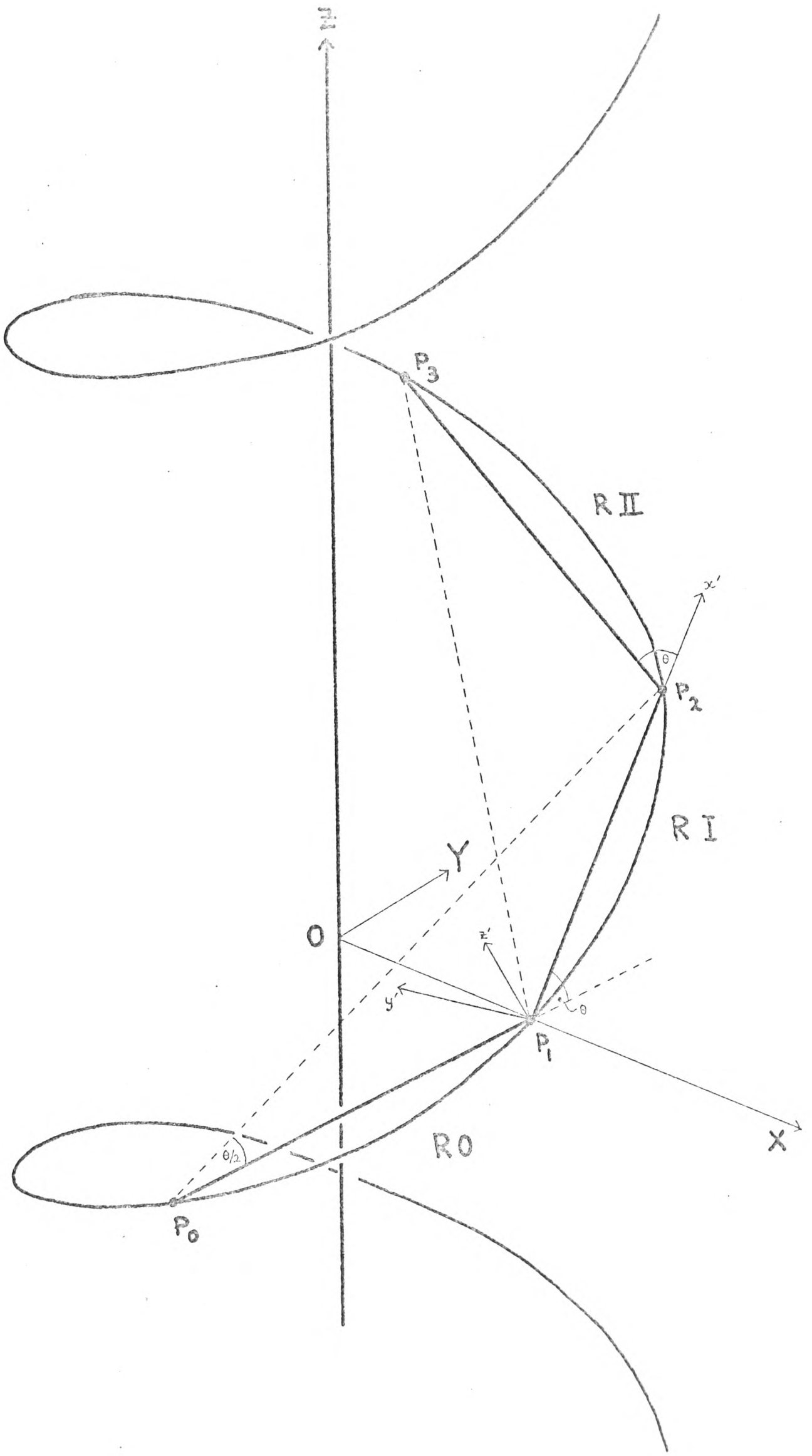


Fig. 31. The geometry of the helix showing sets of coordinate axes used in modelbuild I.

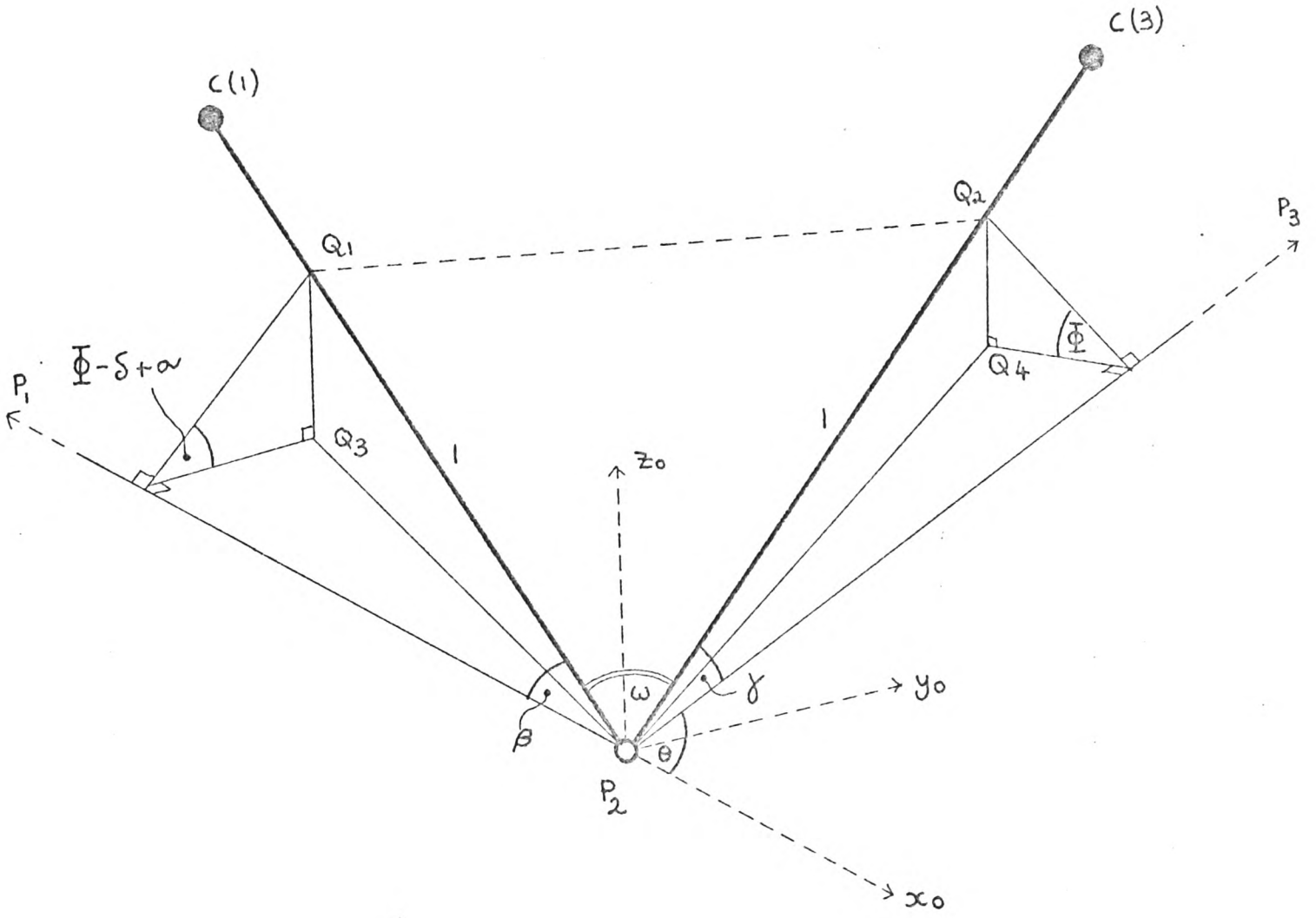


Fig. 32. Details of the geometry at the bridge oxygen atom at P_2.

A set of coordinate axes x_o, y_o, z_o was chosen as shown in Fig.32 and the coordinates of Q_1 and Q_2 were calculated with respect to these.

Q_1 was $-\cos\beta$, $\sin\beta . \cos(\Phi+\alpha)$, $\sin\beta . \sin(\Phi+\alpha)$

Q_2 was $\cos\gamma . \cos\theta - \sin\gamma . \cos\Phi . \sin\theta$, $\cos\gamma . \sin\theta + \sin\gamma . \cos\Phi . \cos\theta$, $\sin\gamma . \sin\Phi$

$(Q_1Q_2)^2$ was calculated in terms of α^*, β, γ, δ and θ and an expression was found for Φ in terms of these angles and ω. This was of the form:-

$$a . \cos^2\Phi + b . \cos\Phi . \sin\Phi + c . \sin\Phi + d . \cos\Phi + e = 0$$

where

$$a = \sin\beta . \sin\gamma . \cos\alpha - \sin\beta . \sin\gamma . \cos\theta . \cos\alpha$$

$$b = \sin\beta . \sin\gamma . \cos\theta . \sin\alpha - \sin\beta . \sin\gamma . \sin\alpha$$

$$c = \sin\beta . \sin\theta . \cos\gamma . \sin\alpha$$

$$d = \sin\gamma . \cos\beta . \sin\theta + \sin\beta . \cos\gamma . \sin\theta . \cos\alpha$$

$$e = \cos\beta . \cos\gamma . \cos\theta + \cos\omega - \sin\gamma . \sin\beta . \cos\alpha$$

This gave a quartic equation in cos Φ

$$a_0 . \cos^4\Phi + 4 . a_1 . \cos^3\Phi + 6 . a_2 . \cos^2\Phi + 4 . a_3 . \cos\Phi + a_4 = 0$$

where

$$a_0 = a^2 + b^2$$
$$a_1 = (2bc - 2ad)/4$$
$$a_2 = (c^2 - b^2 + 2ea + d^2)/6$$
$$a_3 = -(2bc + 2ed)/4$$
$$a_4 = e^2 - c^2$$

This has been expressed in a form suitable for solution by Ferrari's method (66). The reducing cubic of the quartic was:-

$$4\lambda^3 - I . \lambda + J = 0$$

where $I = a_0 . a_4 - 4 . a_1 . a_3 + 3 a_2^2$

and $$J = \begin{vmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{vmatrix}$$

* $\alpha = \omega - \delta$

From the geometry of the system it was apparent that there would be either two real solutions for Φ or no real solutions. When the former was the case it was possible to solve the equation and find the two real roots. This was the case when the discriminant of the quartic equation, $\Delta_q = I^2 - 27J^2$, was negative and the corresponding discriminant of the reducing cubic, $\Delta_c = (27J^2 - I^2)/16$, was positive. If Δ_c was negative then Δ_q was positive and there were no real solutions for Φ indicating that a helix of the required h and n could not be formed for the disaccharide conformation under consideration.

The real root λ_1 of the reducing cubic was found by Tartaglia's method (66). The roots of the quartic equation were then the roots of the two quadratic equations:-

$$a_0 x^2 + 2(a_1 - p_1)x + a_2 + 2\lambda_1 - q_1 = 0$$

and $$a_0 x^2 + 2(a_1 + p_1)x + a_2 + 2\lambda_1 + q_1 = 0$$

where $$p_1 = (a_0.\lambda_1 + a_1^2 - a_0.a_2)^{1/2}$$

and $$q_1 = (2a_1.\lambda_1 + a_1.a_2 - a_0.a_3)/p_1$$

For each solution of Φ the coordinates of disaccharide residue I were found with respect to axes x' , y' , z' (Fig.31) by a rotation around the X3 axis (set along P_1 P_2) of $-\Phi$. The x' axis lay along P_1 P_2 , the y' axis lay in the P_0 P_1 P_2 plane and the z' axis was normal to the P_0 P_1 P_2 plane. The coordinates were then transferred to a set of axes X, Y, Z (the 'helix' axes) with the Z axis 'vertical' lying along the helix axis and with the X axis horizontal lying in the P_0 P_1 P_2 plane/

plane along the bisector of the angle P_0 P_1 P_2. The matrix for the rotational part of this transformation was:-

$$\begin{vmatrix} -\sin(\theta/2) & -\cos(\theta/2) & 0 \\ \cos(\theta/2).\cos\rho & -\sin(\theta/2).\cos\rho & -\sin\rho \\ \cos(\theta/2).\sin\rho & -\sin(\theta/2).\sin\rho & \cos\rho \end{vmatrix}$$

where ρ was the angle between P_1 P_2 and the horizontal.

Having transferred the coordinates of residue I to the helix axes the cartesian coordinates and semi-polar coordinates were output. The coordinates of residue II were generated from the symmetry of the helix and the interatomic contact distances were checked at each of the two linkages. Captions were printed to indicate whether the conformation was or was not allowed at each linkage. As a check on the calculation, the bridge oxygen angle at O(1/3) of the A-B linkage was calculated to make sure that it was the same as the input value.

Left handed helices were treated by taking the mirror image coordinates of the input monosaccharides (by changing the sign of their 'y' coordinates) and by changing the sign of the torsional angles. A right handed helix was built with these mirror image coordinates. The final Y coordinates and the torsional angles ϕ_2 and ψ_2 were again changed in sign. In the program lis$\overset{t}{\wedge}$ed in Appendix I this inversion was done within the program if a negative value was given for the screw symmetry in the input data.

The computer program for modelbuild I was written in Atlas Autocode (Edinburgh University ISO version) and used on a KDF9 computer./

computer. A listing is given in Appendix I of this thesis.

3.9, COORDINATES FOR 3,6-ANHYDROGALACTOSE AND GALACTOSE

A set of coordinates was estimated for 3,6-anhydrogalactose for use with the modelbuilding program and in some of the methods used in the angal crystal structure determination. The set of coordinates was calculated by joining a 'standard' five membered sugar ring C(3), C(4), C(5), C(6), O(3) on to part of a 'standard' six membered sugar ring C(5), O(5), C(1), C(2), C(3). The coordinates were measured from a series of scaled drawings. Bond lengths and bond angles were calculated and a few adjustments were made to some of the atomic positions to give better bond lengths and bond angles. The ring was slightly flattened at C(1) in order to relieve a rather short contact between the hydrogen atom on C(1) and one of the hydrogen atoms on C(6). The coordinates of the model are given in Table 8. A computer program was written to calculate the positions of the hydrogen atoms attached to the carbon atoms such that the angles between the carbon-hydrogen bond and the three other bonds of the carbon atom were all equal. The carbon hydrogen bond length was set at 1.08 Å. The coordinates of the two hydrogen atoms on C(6) could not be calculated by this method and they were estimated from scaled drawings.

A set of coordinates for β-D-galactose was derived from the crystal structure of 6-bromo-6-deoxy-methyl-α-D-galactoside (29). The hydrogen atom placing computer program described above was used to calculate a position for O(1) in the β position, the bond length being set at 1.39 Å. The position of H/

Atom	x	y	z
C(1)	-1.3859	0.0000	0.0000
C(2)	-1.7480	1.2719	0.7710
C(3)	-3.2280	1.2228	1.1386
C(4)	-3.3582	0.0000	2.0488
C(5)	-3.1329	-1.1973	11.1209
C(6)	-3.8690	-0.6703	-0.1308
O(1)	0.0000	0.0000	0.0000
O(2)	-0.9835	1.3248	1.9823
O(3)	-3.8592	0.7499	-0.0533
O(4)	-4.6662	-0.0536	2.6323
O(5)	-1.7480	-1.1481	0.7769
H(1)	-1.8041	-0.0194	-0.9955
H(2)	-1.5443	2.1434	0.1660
H(3)	-3.6327	2.1398	1.5398
H(4)	-2.6030	0.0199	2.8204
H(5)	-3.3816	-2.2461	1.1883
H(6)	-4.9679	-1.0320	0.0090
H(7)	-3.4625	-1.0346	-1.0655

(Coordinates in Angstroms)

Table 8. Coordinates of angal model.

H(1) was set along the direction of the original C(1) - O(1) bond at a distance of 1.08 Å. Other hydrogen atom positions were calculated using the computer program. O(6) was omitted as rotation around the C(5) - C(6) bond was possible.

A computer program was written to transfer the coordinates of any monosaccharide with respect to any set of crystal axes to a set of axes suitable for use in the modelbuild I program or in a steric map calculation for a disaccharide. The numbers of three atoms in the list had to be specified, one to define the origin, a second to define the direction of the x axis and the third to define the xz plane and the positive direction of z.

3.10, CALCULATION OF STERIC MAPS FOR IOTA

A computer program was written to calculate steric maps for disaccharides or for linkages in polysaccharide chains. The input of the monosaccharide coordinates and the first part of the calculation was similar to the first part of the modelbuild I program where, for each combination of ϕ and ψ, the coordinates of the two monosaccharide units adjacant to the linkage were transferred to a common set of axes. The distances between the atoms of adjacant monosaccharide units were calculated. If any of these distances was less than the allowed van der Waals contact distance for that pair then the conformation was disallowed and a 'o' was printed on the map. If there were no such short contacts then a '1' was printed. The angles ϕ and ψ were varied stepwise over 360 degrees usually in ten degree intervals. Two sets of van der Waals contact distances were used, one to

give/

give the 'fully' allowed conformations and one to give the 'marginally' allowed conformations. The values were those used by G.N. Ramachandran, C. Ramakrishnan and V. Sasisekharan (32) and are as follows:-

	Fully allowed	Outer limit (marginally allowed)
C ... C	3.20	3.00
C ... O	2.80	2.70
C ... H	2.40	2.20
O ... O	2.80	2.70
O ... H	2.40	2.20
H ... H	2.00	1.90

(Distances in Angstroms)

In this treatment the atoms are treated as hard spheres and the steric maps are sometimes referred to as hard sphere maps.

Steric maps were calculated for the two iota linkages using both sets of contact distances. The maps are shown in Figs. 33 and 34 where the areas enclosed by the full lines indicate the fully allowed conformations and the areas enclosed by the broken lines indicate the marginally allowed conformations. The atoms included in the calculation of the steric maps were as follows:-

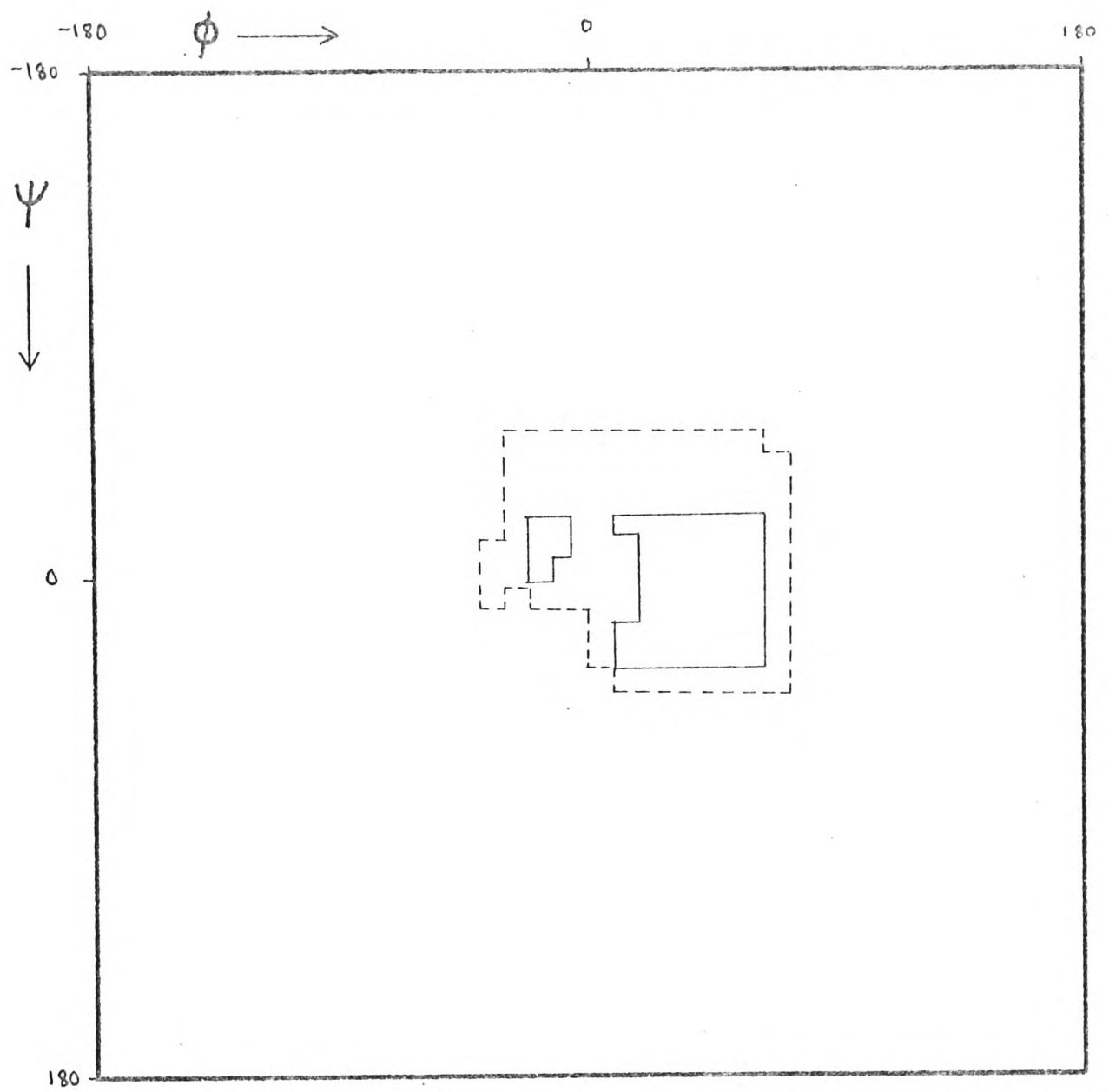
(i) For the galactose-1-β-4-anhydrogalactose linkage

C(1), C(2), O(2), O(5), H(1) and H(2) of galactose

and C(3), C(4), C(5), C(6), O(3), H(3), H(4) and H(6) of anhydrogalactose

The zero positions of ϕ and ψ were defined by C(4) of the galactose and C(1) of the anhydrogalactose respectively. (See Section 3.8 for method of fixing zero positions.)

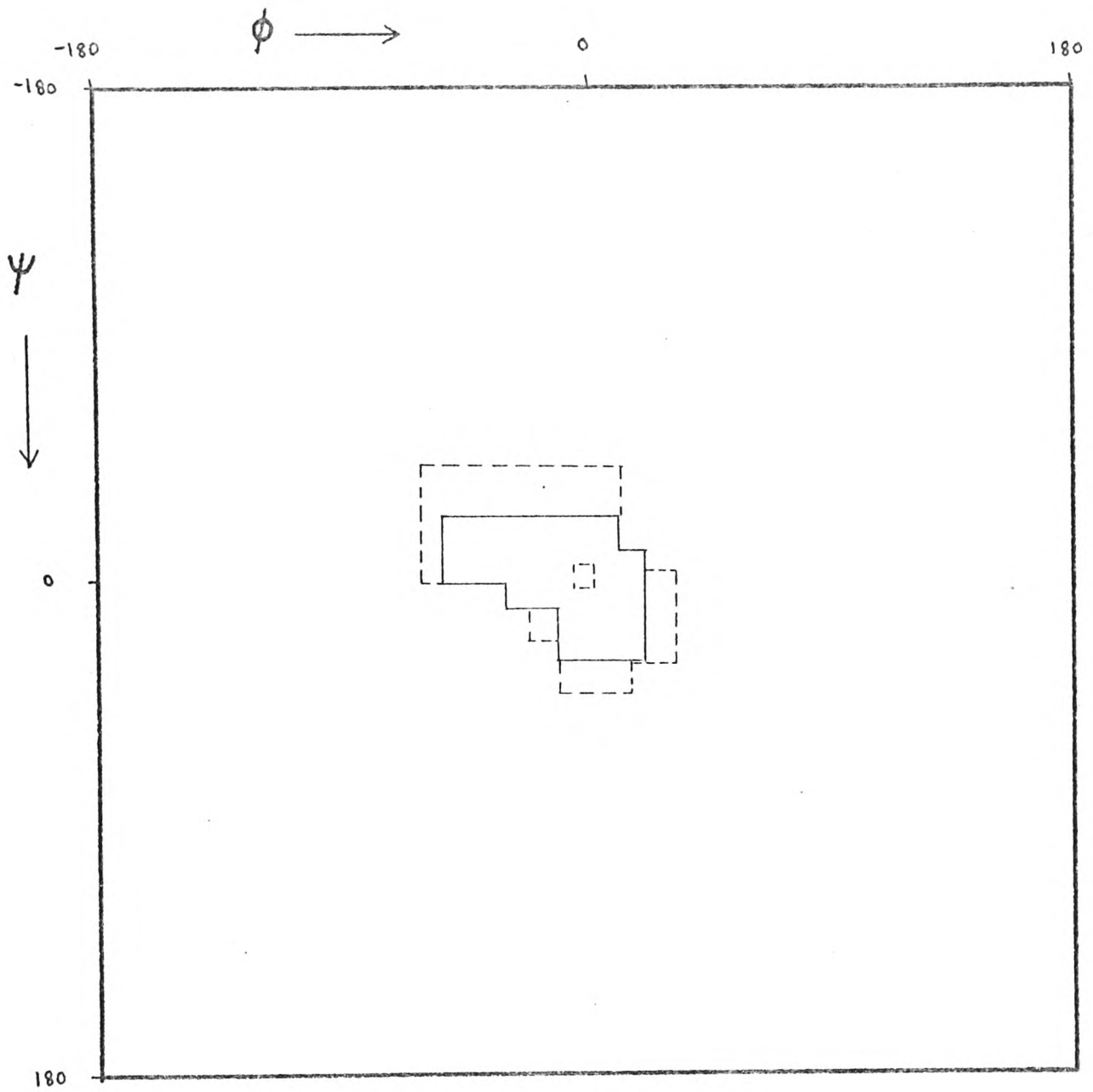
(ii)/



------ marginally allowed area.

——— fully allowed area.

Fig. 33. Steric map of the galactose-1-β-4-3,6-anhydro-galactose linkage of iota.



----- marginally allowed area.
——— fully allowed area.

Fig. 34. Steric map of the 3,6-anhydrogalactose-1-α-3-galactose linkage of iota.

(ii) For the anhydrogalactose -1-α-3- galactose linkage

C(1), C(2), O(2), O(5), H(1), H(2), H(4) and H(7) of anhydrogalactose

and C(2), C(3), C(4), O(2), O(4), H(2), H(3) and H(4) of galactose

The zero positions of ϕ and ψ were defined by C(4) of anhydrogalactose and O(5) of galactose.

3.11, POTENTIAL ENERGY MAPS

In addition to the hard sphere maps, potential energy maps were calculated for the two linkages using the Kitaygorodsky potential energy function (67). The energy for the non-bonded interaction between a pair of atoms is given by

$$E = 3.5\left(-0.04/z^6 + 8600\, e^{-13z}\right) \quad \text{Kcal.mole}^{-1}$$

where z is the distance between a pair of atoms divided by the equilibrium distance for that pair r_o.

The following values of r_o were used (Å):-

	r_o	
C ... C	3.80	
C ... O	3.55	
C ... H	3.15	
O ... O	3.33	
O ... H	3.00	see ref. 68
H ... H	2.60	

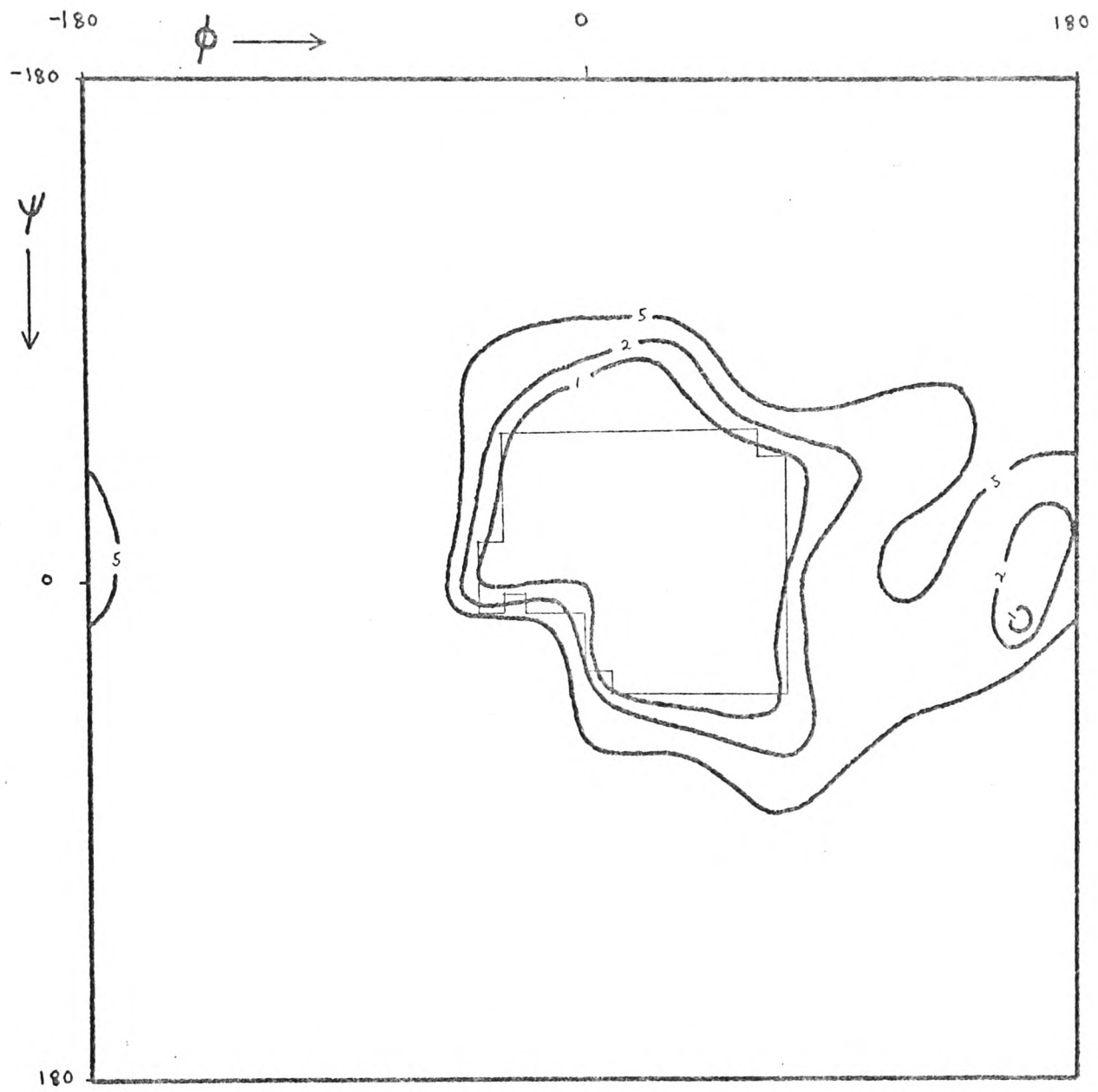
A computer program, similar to the one for the calculation of the hard sphere maps, was used, except that the potential energies of the non-bonded interactions were calculated and summed/

summed instead of making a simple check of van der Waals contact distances. If the distance between a pair of atoms was greater than 5.5 Å then no contribution to the total potential energy was added for that pair and if z was less than 0.2 for any pair the conformation was not considered further. In the latter case the potential energy would be very large and when z becomes small the energy function becomes meaningless and can become negative.

Maps of the Kitaygorodsky potential energy minima are shown in Figs.35 and 36 for the two iota linkages. The same atoms were included in their calculation as for the hard sphere maps. Contours have been drawn at 1, 2 and 5 K cal. $mole^{-1}$ and the areas of marginally allowed conformations are outlined for comparison. In both cases the boundary of the marginally allowed area is followed fairly closely by the 1 Kcal. $mole^{-1}$ contour. In the case of the galactose-1-β-4-anhydrogalactose linkage there is a small second potential energy minimum. A very small allowed area was found in this region in some hard sphere calculations carried out by Dr. D.A. Rees for this linkage though it did not appear in the maps calculated by the author. It would probably be unwise to attach too much importance to the magnitude of the potential energies calculated, especially as it depends so critically on the equilibrium distances used. The potential energy maps could however be useful in showing up potential energy minima which just fail to appear on the hard sphere maps.

3.12, APPLICATION OF MODELBUILD I TO IOTA

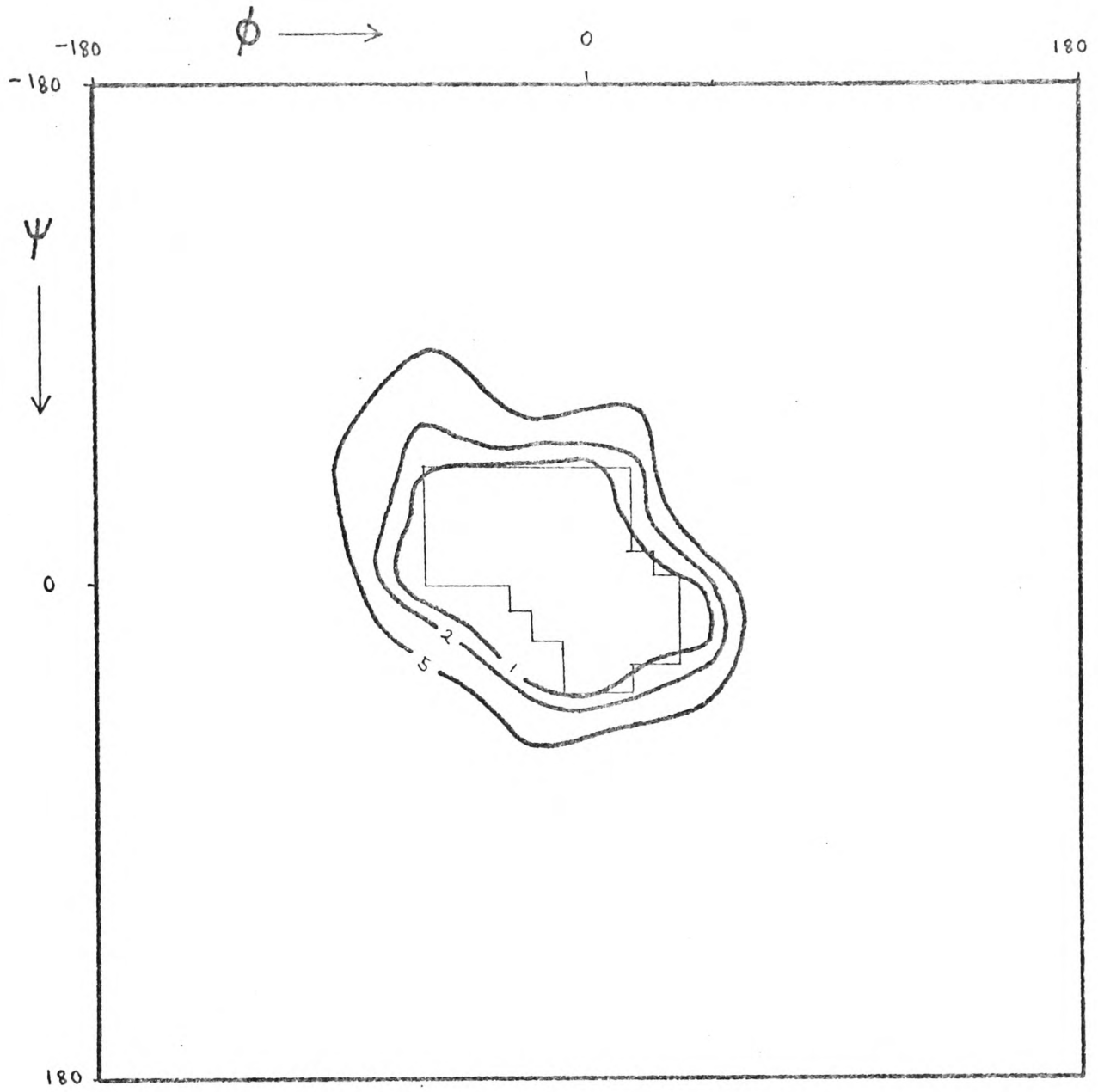
In/



potential energy contours.

marginally allowed area.

Fig. 35. Potential energy map of the galactose-1-β-4-3,6-anhydrogalactose linkage of iota. (Contours at 1,2 and 5 Kcal.mole^{-1}. Marginally allowed area shown for comparison.)



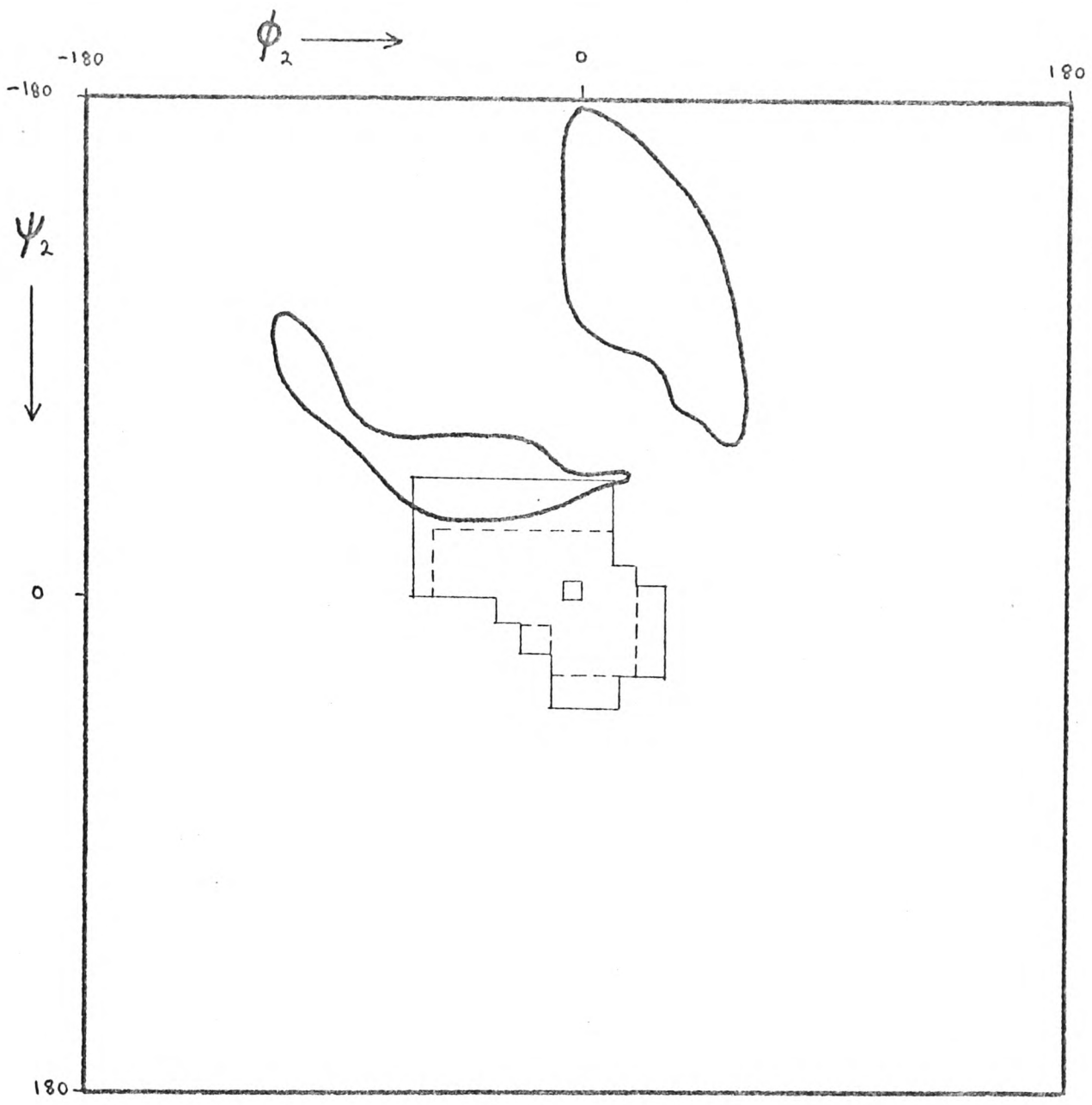
potential energy contours.

marginally allowed area.

Fig. 36. Potential energy map of the 3,6-anhydrogalactose-1-α-3-galactose linkage of iota. (Contours at 1,2 and 5 Kcal.mole^{-1}. Marginally allowed area shown for comparison.)

In its original version the modelbuild I program was written to deal with right handed threefold helices only and without any checks of steric restrictions. The first calculations on iota were carried out with this version of the program. The angles ϕ_1 and ψ_1 at the B-A linkage (galactose-1-β-4-anhydrogalactose) were varied over the ranges -40 degrees to 70 degrees and -50 degrees to 40 degrees respectively in ten degree intervals. These ranges covered the marginally allowed area for the B-A linkage. Hydrogen atoms were not included in the calculation as no inter-atomic contacts were being checked. The two sulphate groups on O(4) of galactose and O(2) of anhydrogalactose were omitted from the calculation as their exact positions were uncertain due to rotational freedom around the single bonds attaching them to the rings. For the same reason O(6) of galactose was also omitted. For most of the conformations covered, two solutions were found for ϕ_2 and ψ_2. These solutions fell in two areas of the anhydrogalactose-1-α-3-galactose steric map as indicated by the full lines in Fig.37. The marginally allowed area is also shown (narrow line). This plot shows that one of the sets of solutions overlaps the marginally allowed area giving a series of possible solutions for the conformation of the single polysaccharide chain. The second set of solutions gave no possible conformations.

The threefold helix program was extended in order to test the feasibility of formation of the symmetry related double helix. The coordinates of residues O to V were generated. The coordinates of a single disaccharide residue of the second chain were/



area of solutions of ϕ_2 and ψ_2.

——— marginally allowed area.

----- fully allowed area.

Fig. 37. The solutions for ϕ_2 and ψ_2 at the A-B linkage calculated for the allowed conformations at the B-A linkage and plotted on the steric map of the A-B linkage.

were calculated such that this residue, residue S, was displaced by half a helix repeat distance, parallel to the helix axis, from residue I. Hydrogen atoms, except those on the hydroxyl groups, were included. The sulphate groups and O(6) of galactose were omitted. The contacts between the atoms of residue S and those of residues O-V were examined. Contacts at each of the two linkages within the chain were also examined and captions were printed to indicate whether the conformation at a linkage was or was not allowed and whether the conformation with a second chain fitted in was or was not allowed. Checking the inter-atomic contacts as part of the calculation made it unnecessary to plot out the results as was done in the first calculation and possible solutions were readily found by looking through the computer output. The original method does have some advantages as 'near misses' can be spotted. Using the second version of the program 25 possible conformations were found for the single chain forming a right handed three-fold helix of pitch 26.0 Å. Of these seventeen allowed the formation of the symmetry related double helix.

From the results plotted in Fig.37 it was apparent that the overlapping region did not extend into the fully allowed region of conformations so that no solution would be expected if the fully allowed contact distances were used. The contact distances were therefore increased by 0.1 Å only, in order to narrow down the number of 'possible' conformations. With this set of contact distances only one possible conformation was found for the double helix. This conformation was designated model I. The coordinates of model I are listed in Table 9.

The/

Atom	Phi °	R(Å)	Z (Å)	
C(1)	66.943	2.843	7.844	
C(2)	78.752	3.395	6.567	
C(3)	61.255	3.466	5.457	
C(4)	43.598	2.208	5.264	
C(5)	22.988	2.417	6.536	
C(6)	30.582	3.893	6.740	
O(1)	90.000	2.680	8.667	
O(2)	102.313	2.886	6.166	3,6-anhydrogalactose
O(3)	47.295	4.328	6.099	
O(4)	25.677	2.609	4.109	
O(5)	43.509	1.771	7.574	
H(1)	57.036	3.655	8.289	
H(2)	82.250	4.435	6.741	
H(3)	66.971	3.937	4.557	
H(4)	64.420	1.328	5.173	
H(5)	359.695	2.660	6.780	
H(6)	19.982	4.526	6.143	
H(7)	30.571	4.209	7.778	
C(1)	357.458	2.003	3.539	
C(2)	358.867	2.247	2.037	
C(3)	324.552	2.434	1.392	
C(4)	319.490	3.740	2.059	
C(5)	315.973	3.487	3.540	
C(6)	313.817	4.755	4.318	
O(1)	25.677	2.609	4.109	galactose
O(2)	33.682	1.656	1.409	
O(3)	330.000	2.680	0.000	
O(4)	332.146	4.791	1.947	
O(5)	337.805	2.991	4.127	
H(1)	344.027	0.986	3.687	
H(2)	4.147	3.282	1.855	
H(3)	299.313	1.924	1.506	
H(4)	307.382	4.280	1.656	
H(5)	298.837	3.020	3.654	
S	107.0	4.30	5.80	3,6-anhydrogalactose
O(a)	100.0	4.10	4.45	
O(b)	95.0	5.60	6.35	(or z-0.5Å)
O(c)	118.0	4.90	5.75	
O(6)	316.0	4.50	6.00	
S	322.0	5.85	1.60	galactose
O(a)	315.0	5.50	0.10	
O(b)	330.0	6.90	1.30	
O(c)	315.0	6.50	2.50	
Mg	337.0	5.50	-0.38	cation (with second set of coordinates for sulphate group of 3,6-anhydro-galactose. i.e. with the z coordinates 0.5Å less than those quoted in the table.)

Table 9, Semi-polar coordinates of iota model I.

The investigation of left handed helices was carried out by building a right handed helix for the mirror image configurations of the monosaccharides. Using the marginally allowed contact distances 15 possible conformations were found for the single chain but no solutions were found for a left handed double helix.

The modelbuild I calculations showed that a threefold right handed double helix, with a pitch of 26.0 Å, was sterically feasible for iota. The conformation would probably be close to that of model I though with a possible variation of up to about 25 degrees in the four torsional angles. A threefold right handed or left handed single helix of pitch 13.0 Å was still formally a possibility and possible conformations for such a helix were found by Dr. D.A. Rees using modelbuild II. The diameters of the single helices were greater than that of the double helix and it would be difficult to fit them in to the proposed packing arrangement. The results of intensity calculations also favoured the double helix model (see Section 5.2).

3.13, NUMBERS OF POSSIBLE CONFORMATIONS

Table 10 gives an estimate of the number of conformations involved for iota under various conditions. If all four torsional angles are varied over 360 degrees in ten degree intervals then the total number of conformations is about 1,680,000. If the conformations are restricted to those which are sterically feasible there are about 6,500 possible conformations and if the condition is imposed that a helix of given h

and/

CONDITIONS	NUMBER OF CONFORMATIONS	SENSE OF HELICES
ALL CONFORMATIONS (Vary ϕ_1, ψ_1, ϕ_2 & ψ_2)	1,680,000	R.H. & L.H.
STERIC RESTRICTIONS	6,500	R.H. & L.H.
h & n RESTRICTIONS	5,000	R.H. & L.H.
h & n AND STERIC RESTRICTIONS	25 15	R.H. L.H.
SYMMETRY RELATED DOUBLE HELIX	17 0	R.H. L.H.

Table 10. Numbers of conformations for iota under various conditions.

and n must be formed there are about 5,000 possible conformations. When both conditions are imposed simultaneously as in modelbuild I there are forty conformations, 25 right handed helices and 15 left handed helices. 17 of the right handed helices allowed the formation of the symmetry related double helix. This is only 0.001% of the total number of conformations if no restrictions are imposed.

3.14. PHYSICAL MODELS

Having selected a single conformation of the polysaccharide chain (model I) as probably being a good approximation to the true conformation, models were built using Beevers models (70). Making use of the coordinates which had been calculated with respect to the helix axes, the bridge oxygen atoms were specially drilled and attached to a central perspex rod by steel rods of the required lengths. The monosaccharide units were then inserted between the bridge oxygen atoms. Coordinates of the atoms were checked to make sure that they were consistent with the calculated values.

The positions of the sulphate groups occurred on the outside of the helix and both of them appeared to have some degree of freedom. As they had been omitted from the calculation of the chain conformation, the occurrence of sensible positions for these groups helped to confirm the reasonableness of the model proposed. A possible hydrogen bond perpendicular to the helix axis was also noticed from the model. This was between O(6) of galactose in one chain and O(2) of galactose in the second chain.

This/

This involves the only two free hydroxyl groups in the structure, again confirming the reasonableness of the predicted conformation. From polarised infra-red radiation studies on stretched fibres of iota and kappa evidence of a hydrogen bond, of length 2.8 - 2.9 Å and perpendicular to the helix axis, was obtained by F.B. Williamson (unpublished but see ref.52). Photographs of a single helix and a double helix are shown in Fig.50.

A space filling model of iota was made by Mr. F.B. Williamson using CPK space filling models. A photograph of this is shown in Fig.50.

SECTION IV: CYLINDRICALLY AVERAGED PATTERSON FUNCTIONS

4.1, CYLINDRICALLY AVERAGED PATTERSON FUNCTIONS

Cylindrically averaged Patterson functions (71) were calculated from the observed intensity traces for several of the iota salts. The expression for the Patterson function is:-

$$\phi(z,r) = \sum_{l} \phi_l(r) \cdot \cos(2.\pi.l.z)$$

where $$\phi_l(r) = \int_{R=0}^{\infty} I(l,R) \cdot J_0(2.\pi.R.r).R.dR.$$

r = radius in Å

z = fractional coordinate along fibre axis

l = index of the layer line

$I(l,R)$ = the intensity at radius R ($Å^{-1}$) in reciprocal space on the lth layer line.

In practice $\phi_l(R)$ could only be approximately calculated as the observed range of R was restricted. The integral was replaced by a sum and the expression calculated was:-

$$\phi_l(r) = \sum_{R=0}^{R_{max}} I(l,R) \cdot J_0(2.\pi.R.r).R.\Delta R.$$

A computer program was written to calculate this function. Values of $J_0(X)$, a zero order Bessel function of argument X, were tabulated at intervals of 0.1 in X from X = 0 to X = 17.5 (72). Values of $J_0(X)$ with larger arguments were calculated from the approximate expression (59).

$$J_0(X) \simeq (\sin X + \cos X)/\sqrt{(\pi.X)}$$

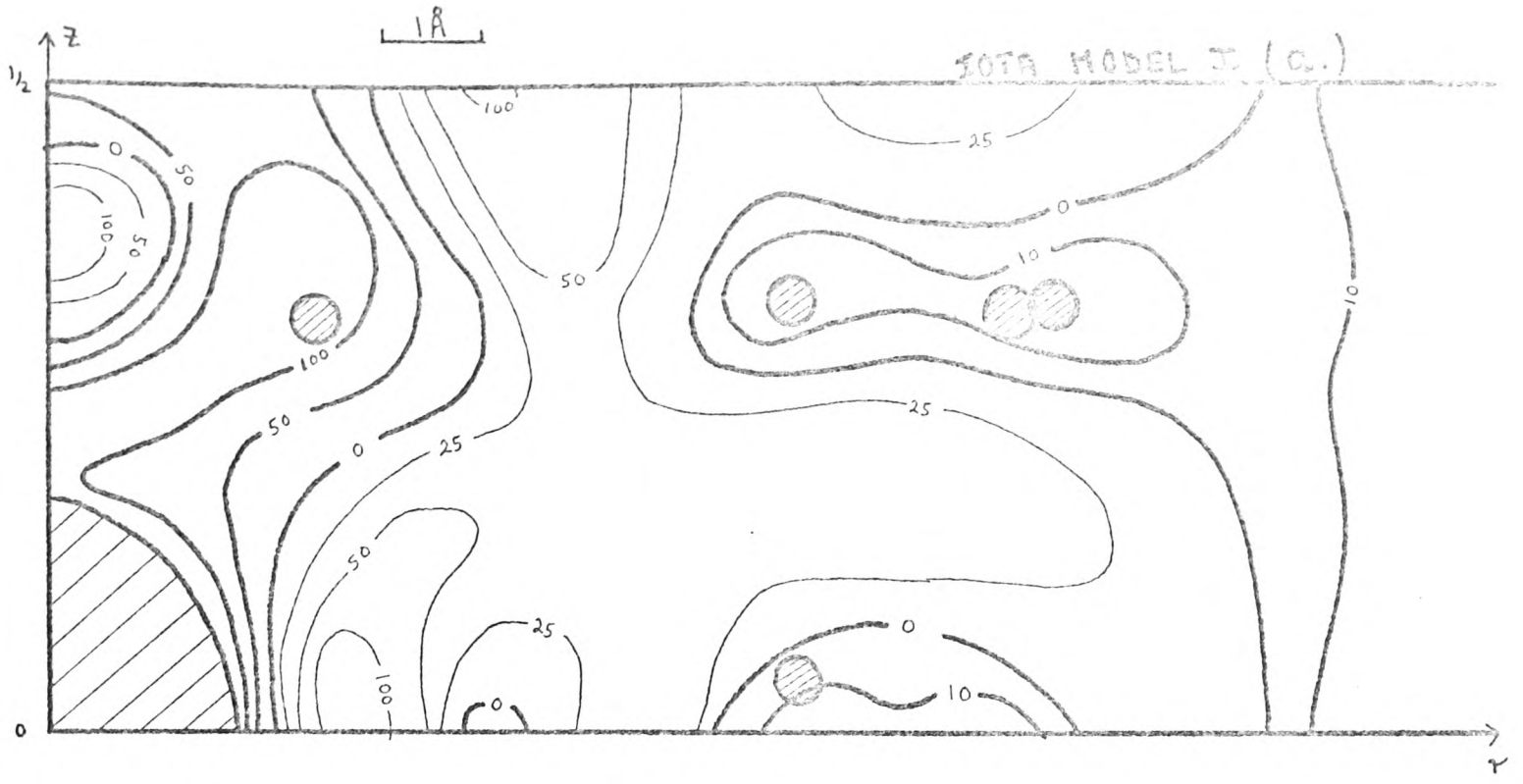
The/

The input intensities were measured at intervals of 0.01 in ξ (= R. λ), the first measurement being at ξ = 0.005. A version was also written to take intensities at intervals of 0.02 in ξ starting at ξ = 0. This latter version was written so that a series of intensities calculated by Dr. M.M. Harding's cylindrically averaged Fourier transform program (Section 5.1) could be directly input. The Patterson functions were calculated at intervals of $^1/30$ in z and normally at intervals of 0.5 A in r .

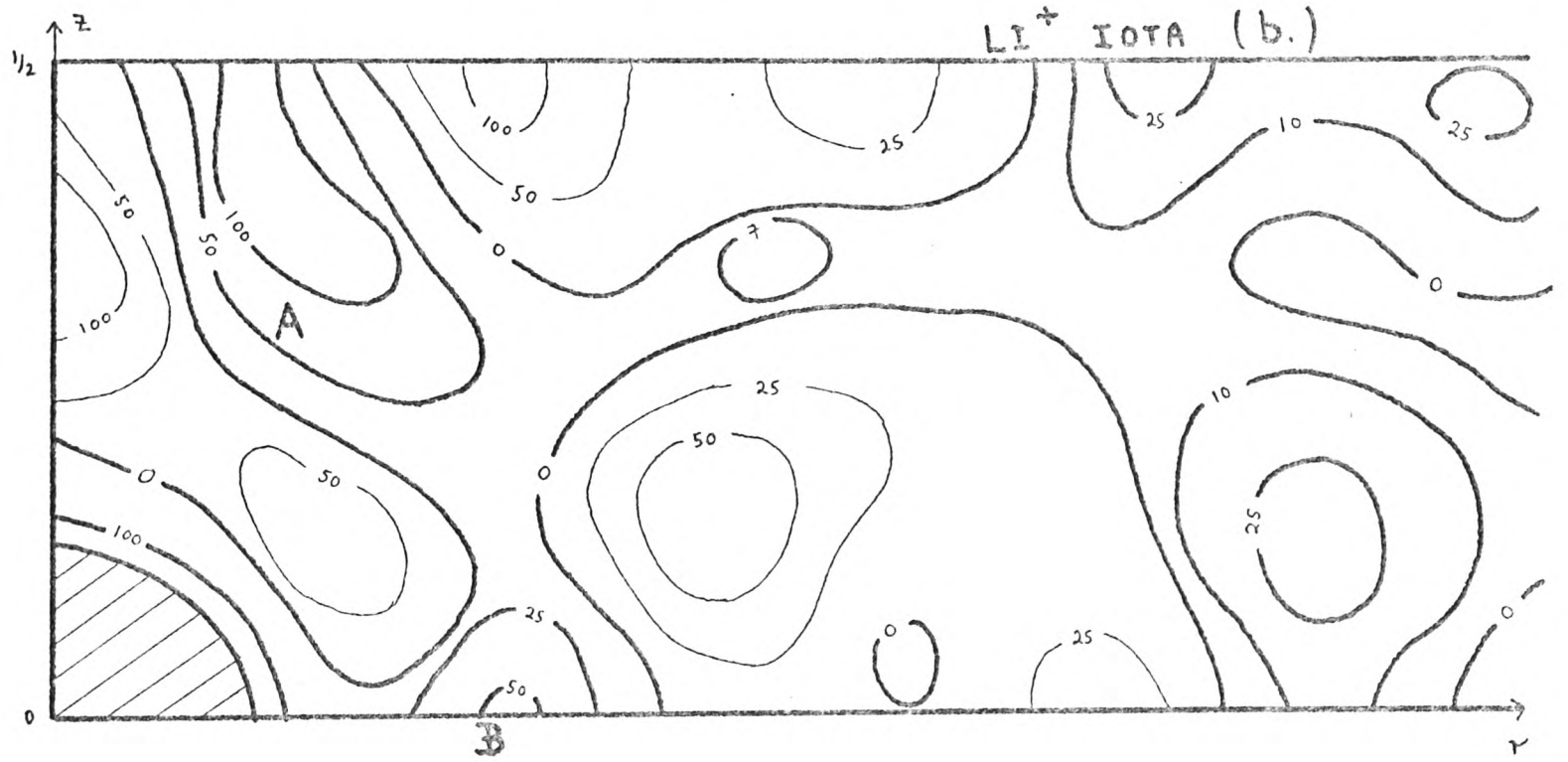
In all the Patterson functions that were calculated, there were severe diffraction effects round the peaks especially at the origin, probably due to the restricted quantity of data available. The heights of the peaks fell off rapidly with increasing r and it was thus not suitable to draw contours on the Patterson maps at evenly spaced intervals. In the maps that are included in this thesis, contours have been drawn to indicate the main features of the Patterson functions and their relative values have been marked.

4.2, IOTA MODEL I PATTERSON FUNCTION

For purposes of comparison, a cylindrically averaged Patterson function was computed from the calculated cylindrically averaged Fourier transform of iota model I. The quantity of data used was comparable to that available from a typical fibre photograph. The function is shown in Fig.38a along with the positions of the sulphate -sulphate vectors of the model. These vectors show up quite clearly and, indeed, the main features of the function can be accounted for in terms of these vectors alone./



sulphate-sulphate vectors of model I.



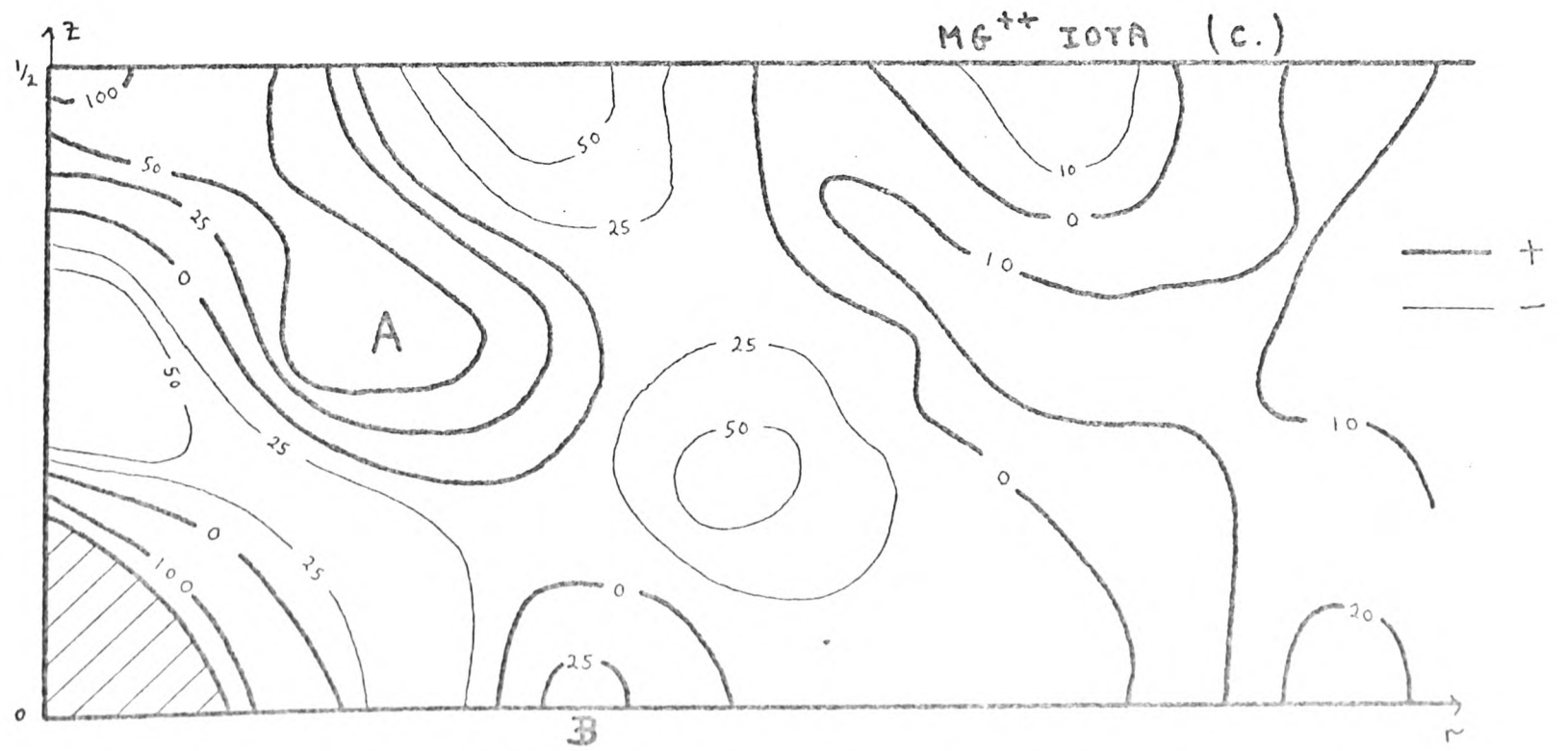


Fig. 38. Cylindrically averaged Patterson functions for model I (a), lithium iota (b) and magnesium iota (c).

alone. This suggests that it should be possible to locate the sulphate -sulphate vectors in the Patterson function of the lithium salt where the cation has very small X-ray scattering power and in that of the magnesium salt where the cation is also relatively light. The lithium and magnesium salt Patterson functions are shown in Fig.38b,c. They show a distinct similarity to the function calculated for model I, though there are extra vectors present, for example the peak B at Z = 0 and $r \simeq 5.5$ Å. This is reasonable as sulphate -sulphate vectors between adjacant double helices will be present.

As not all the sulphate -sulphate vectors within a double helix fall on positive regions of the Patterson functions of the lithium and magnesium salts, any interpretation of the Patterson function in terms of inter double-helix vectors (assuming that model I is approximately correct) would probably have some of the sulphate -sulphate vectors falling on negative regions of the Patterson functions. The differences between the lithium and magnesium salt Patterson functions could be due to slightly different packing arrangements or merely due to poor data.

4.3. POTASSIUM SALT PATTERSON FUNCTION AND DIFFERENCE PATTERSON FUNCTIONS

A Patterson function was calculated for the potassium salt and difference Patterson functions (73) were calculated for the potassium salt and the rubidium salt. One of the problems in calculating difference Patterson functions was scaling the intensities to those of the lithium salt. This was done very approximately/

approximately such that

$$\frac{\Sigma I \text{ at intervals of } 0{\cdot}01 \text{ in } \S \text{ for } C^+ \text{salt}}{\Sigma I \text{ at intervals of } 0{\cdot}01 \text{ in } \S \text{ for } Li^+ \text{ salt}} = \frac{\Sigma f_i^2 \; C^+ \text{ salt}}{\Sigma f_i^2 \; Li^+ \text{ salt}}$$

where $C^+ = K^+$ or Rb^+

The coefficients used for the difference Patterson functions were $|F^2(C^+) - F^2(Li^+)|$ though it might have been better to use $\{\sqrt{(F^2(C^+))} - \sqrt{(F^2(Li^+)}\}^2$ (74).

The potassium salt Patterson function is shown in Fig.39a and the difference Patterson functions in Fig.39b,c. These Patterson functions show a similar vector pattern to that of the lithium salt Patterson function suggesting that the vector pattern of the cations is similar to that of the sulphate groups. The peaks marked A and B are prominent features of all these functions.

4.4, PACKING ARRANGEMENTS

Possible vector patterns, arising from inter double-helical sulphate-sulphate vectors, were considered in terms of possible packing arrangements of the double helices based on space group R3 where the double helices are packed round alternating three-fold and right handed threefold screw axes. The conformation of the double-helix considered was that of model I. Taking orientations of the double helices at 15 degree intervals round the helix axis gave eight possible orientations. The first four of these at 0, 15, 30 and 45 degrees gave the arrangements of sulphate groups Ia, Ib, Ic and Id (Fig.40) around the threefold rotation axis. The orientations at 60, 75, 90 and 105 degrees gave/

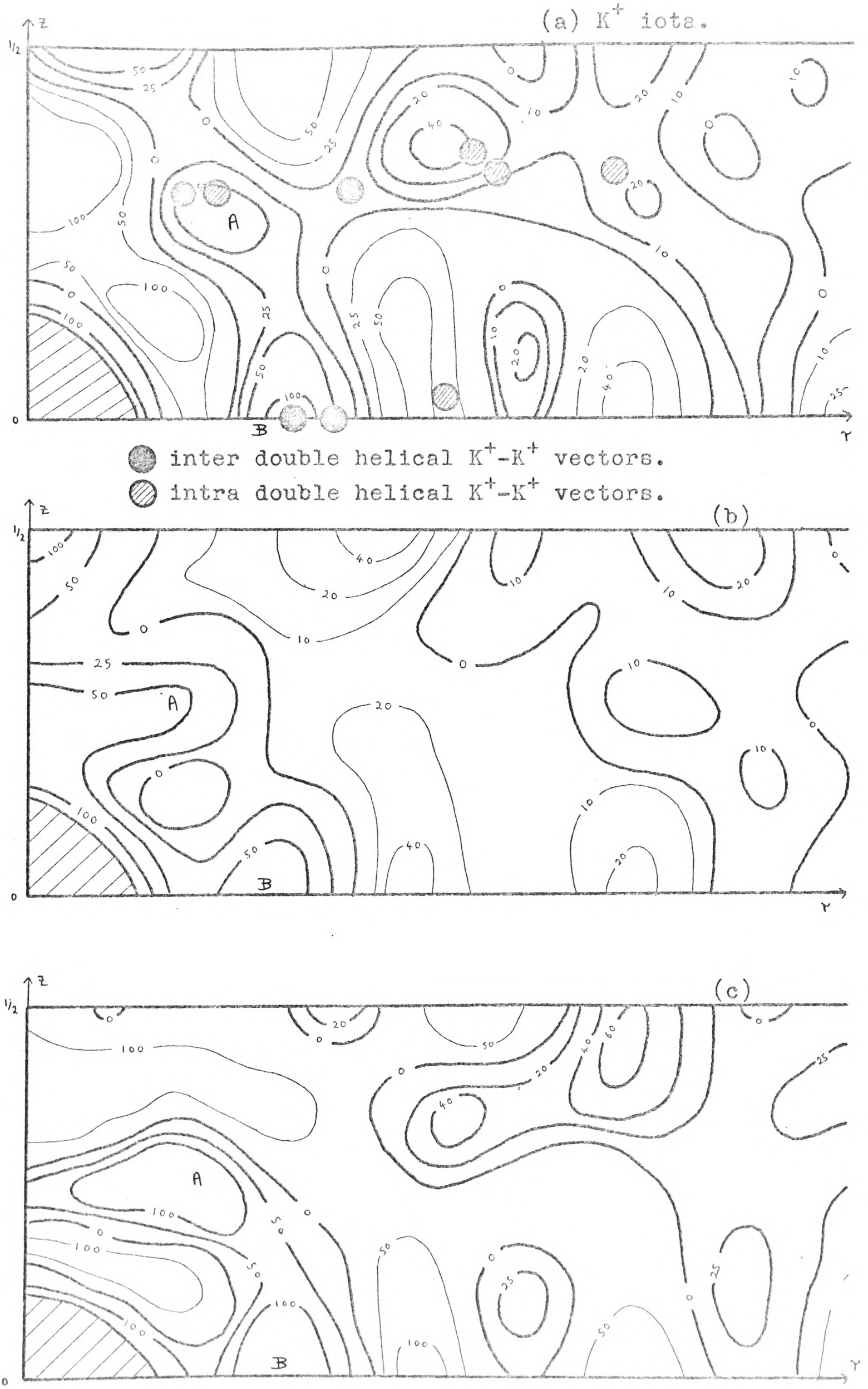


Fig. 39. Cylindrically averaged Patterson function for potassium iota (a) and difference Patterson functions for potassium iota (b) and rubidium iota (c).

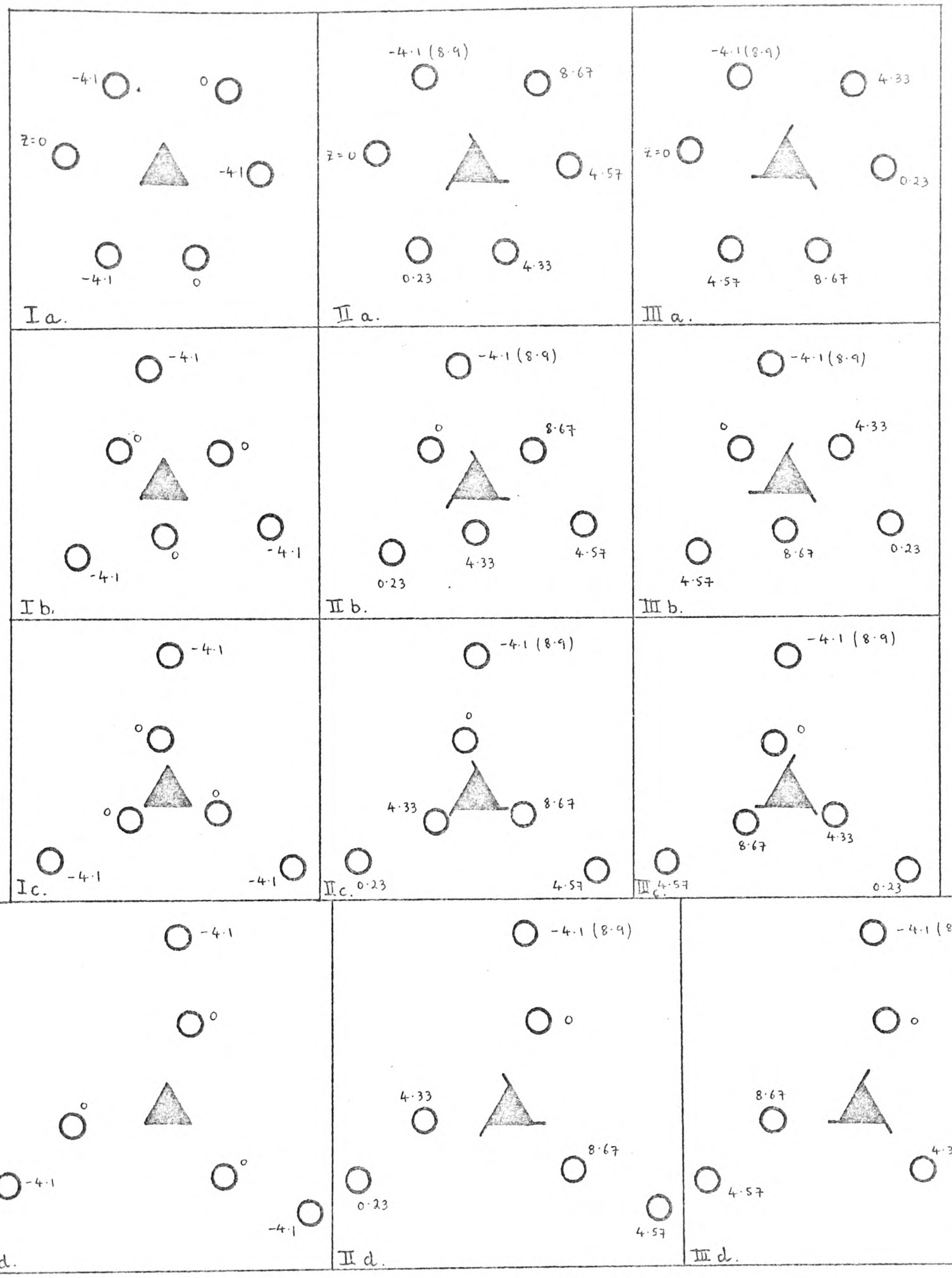


Fig. 40. Possible arrangements of sulphate groups around a three-fold rotation axis, a 3_1 screw axis and a 3_2 screw axis. Each diagram shows two sulphate groups from each of three double helices.

gave the arrangements of sulphate groups IIa, IIb, IIc and IId around the threefold screw axis. The zero position from which these angles were measured had no special significance. Four of the arrangements Ib, Ic, IIa and IIb could be rejected on the grounds that some of the sulphate groups came too close to each other. The other four gave reasonable spacing of the sulphate groups and their vector patterns, superimposed on the lithium salt Patterson function, are shown in Figs.41 and 42. All of these fit the lithium salt Patterson function fairly well though IId does not account for peak B. There are of course many more inter double-helical sulphate-sulphate vectors though these are mostly at larger values of r .

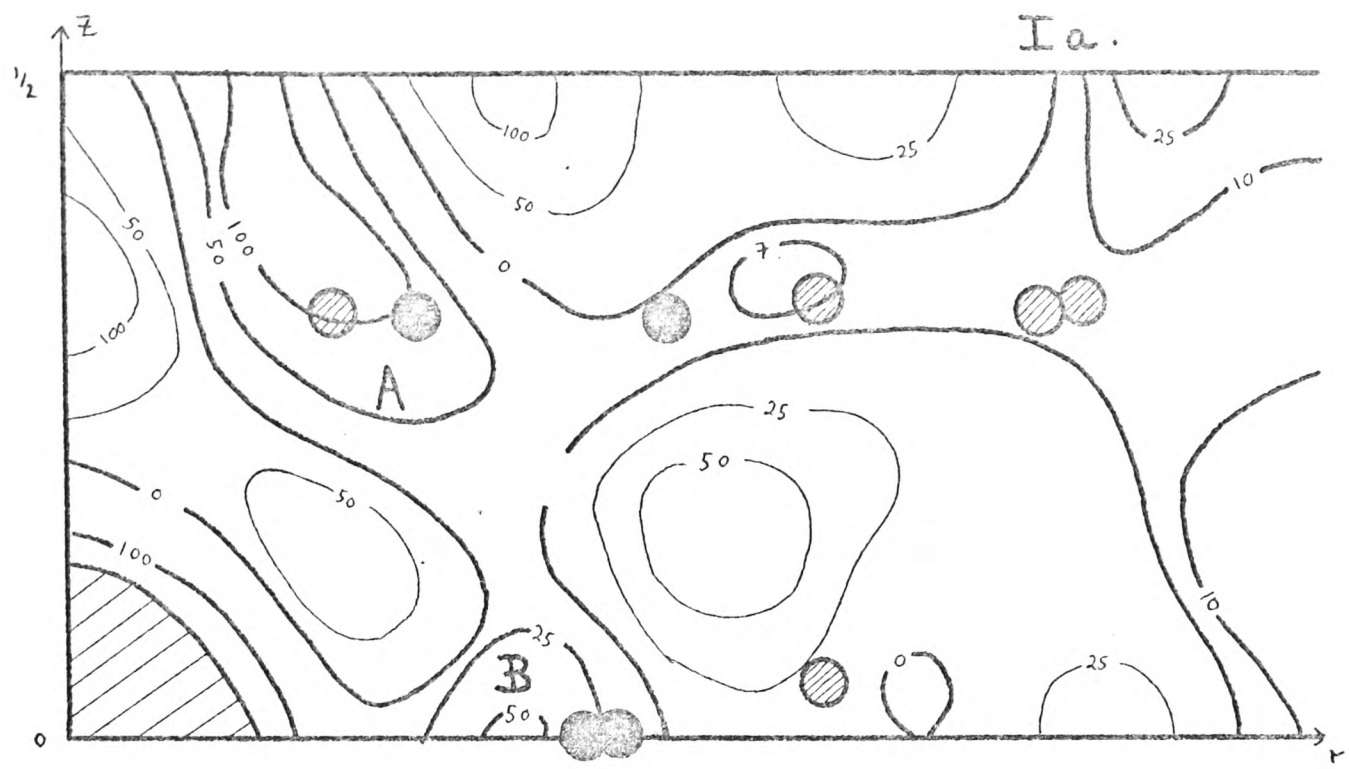
Some packing arrangements, IIIa, IIIb, IIIc and IIId (Fig.40), around a 3_2 screw axis were considered. They would seem to be quite reasonable as far as the spacing of the sulphate group is concerned but were rejected on the grounds that systematic packing round 3_2 screw axes would give space group $P3_2$ [+] and a trigonal unit cell of side ~13 Å.

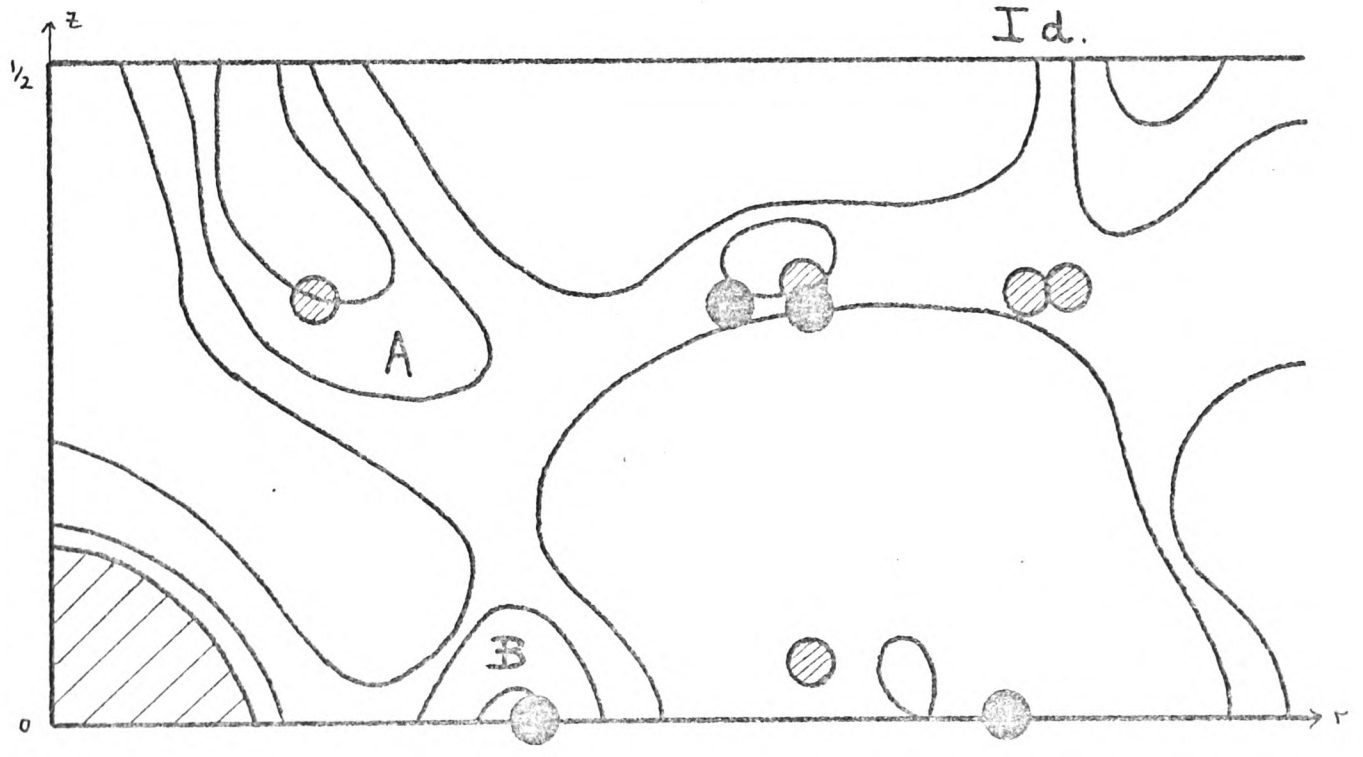
This analysis was not carried further as it seemed unlikely that the questions of orientation and packing could be solved from the Patterson functions of the light cation salts.

4.5, A SOLUTION INCLUDING CATION POSITIONS

One arrangement including cations was proposed. This accounted quite well for the main peaks of the Patterson functions and/

+ See Note on page 70

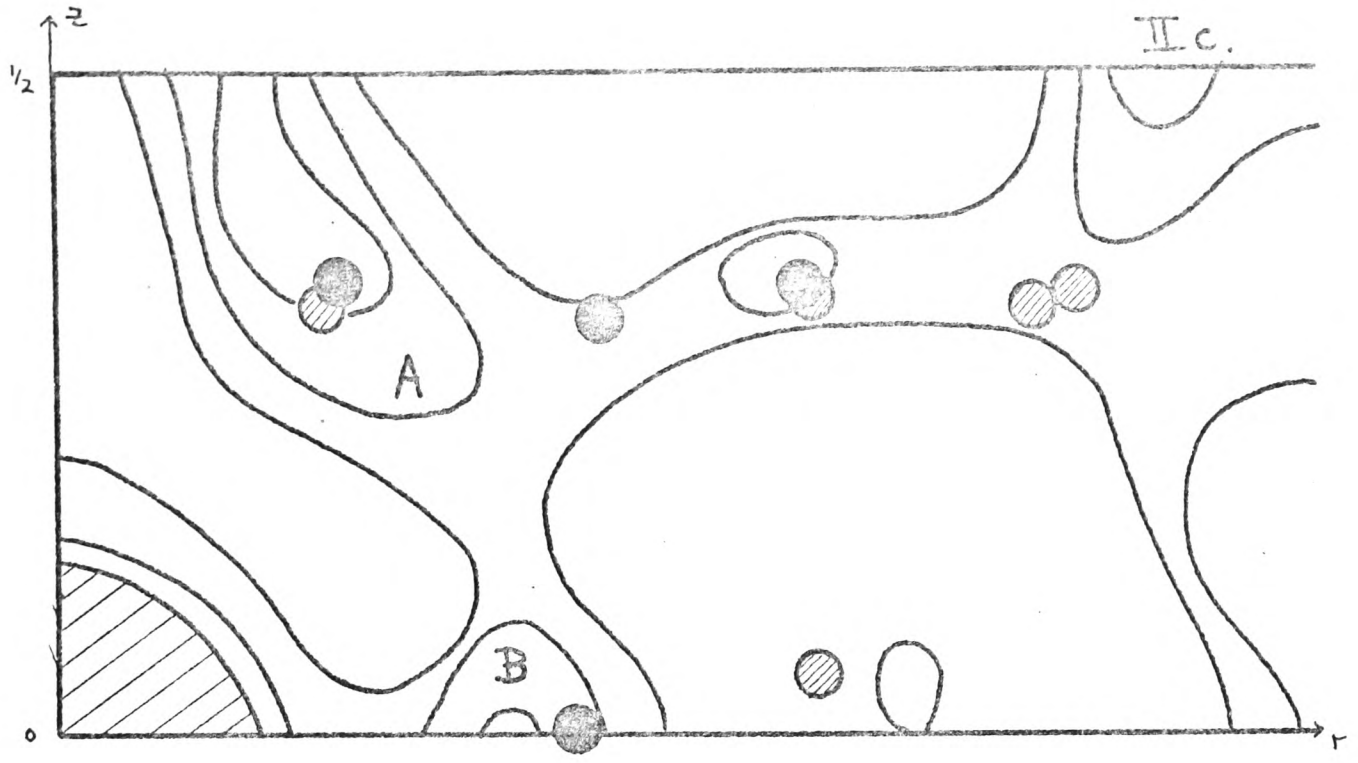


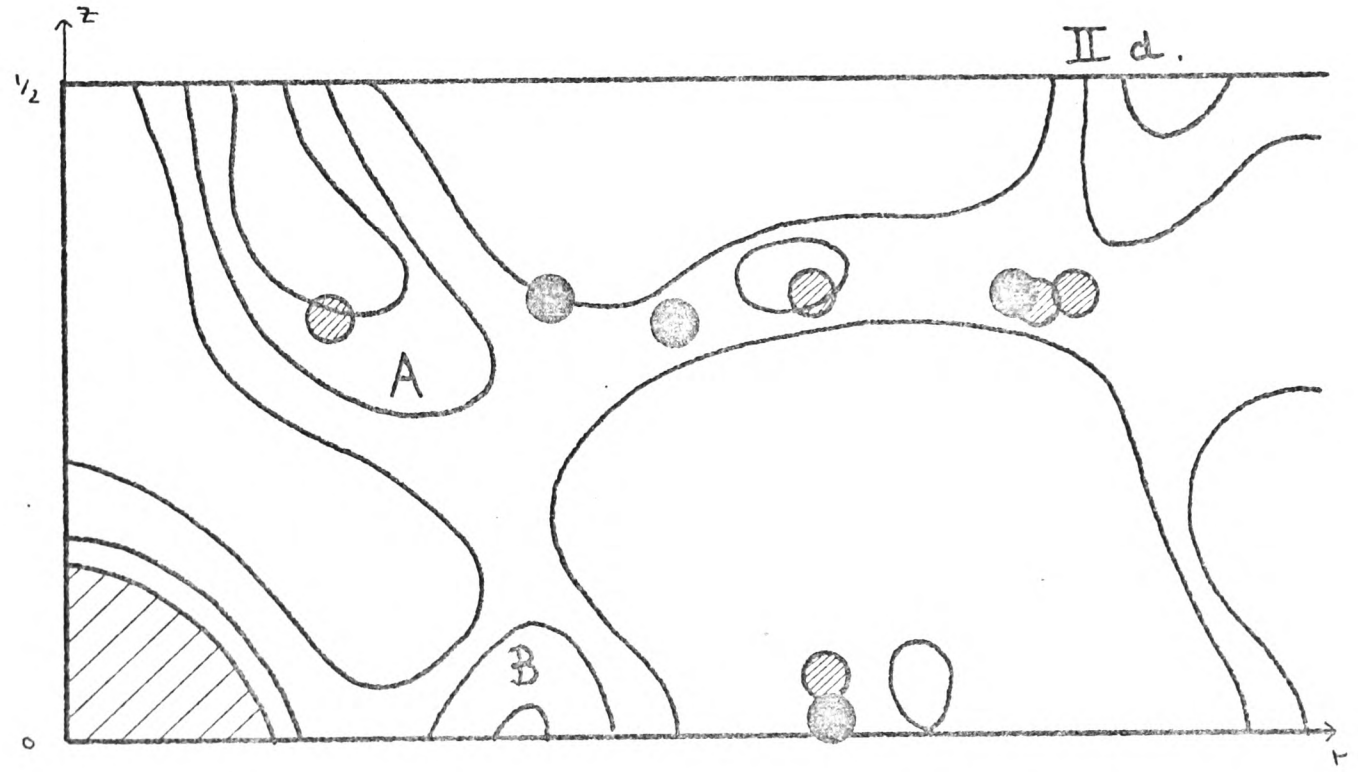


inter double helical

intra double helical

Fig. 41. Sulphate-sulphate vectors from the arrangements Ia and Id superimposed on the lithium iota cylindrically averaged Patterson function.





inter double helical

intra double helical

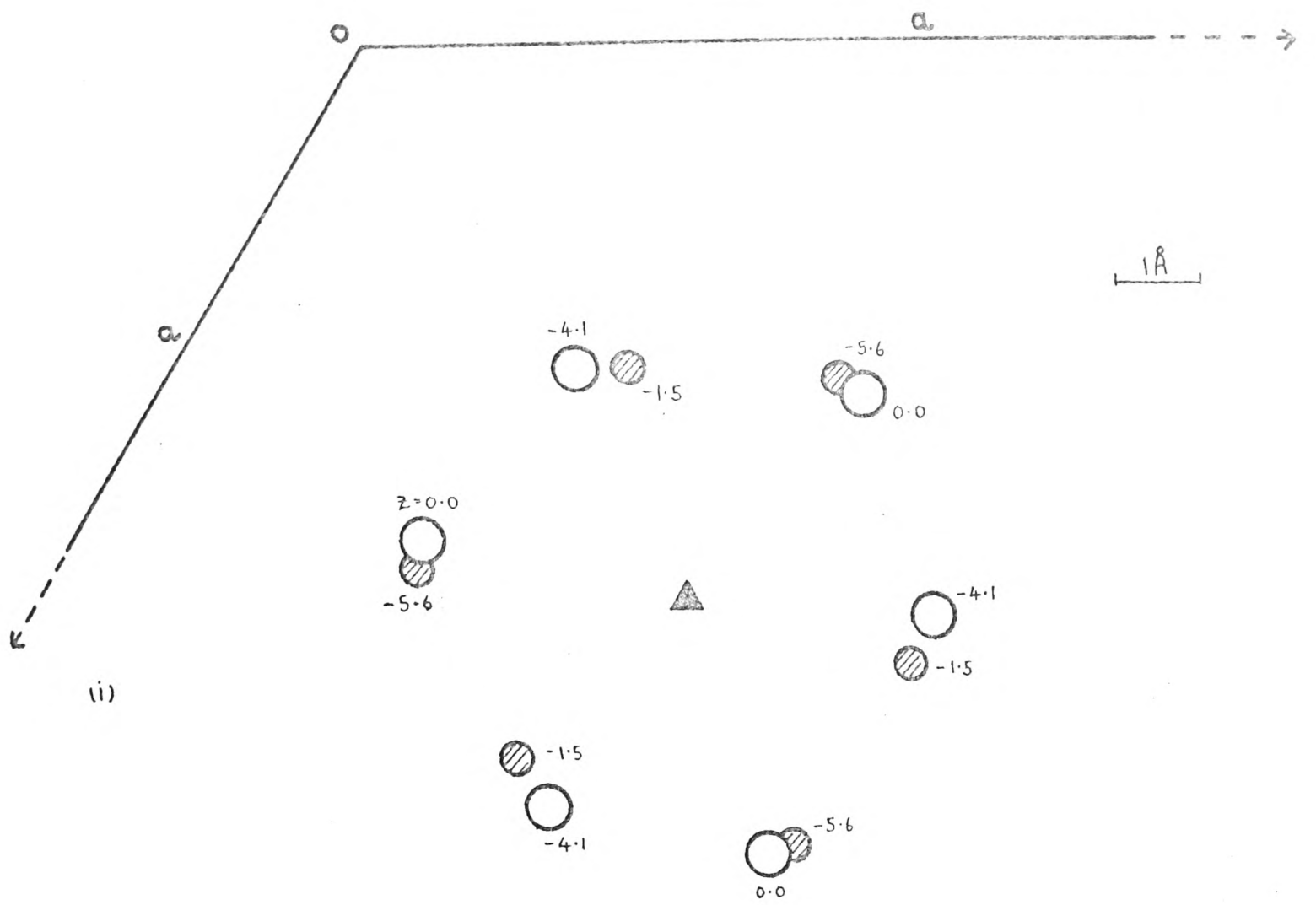
Fig. 42. Sulphate-sulphate vectors from the arrangements IIc and IId superimposed on the lithium iota cylindrically averaged Patterson function.

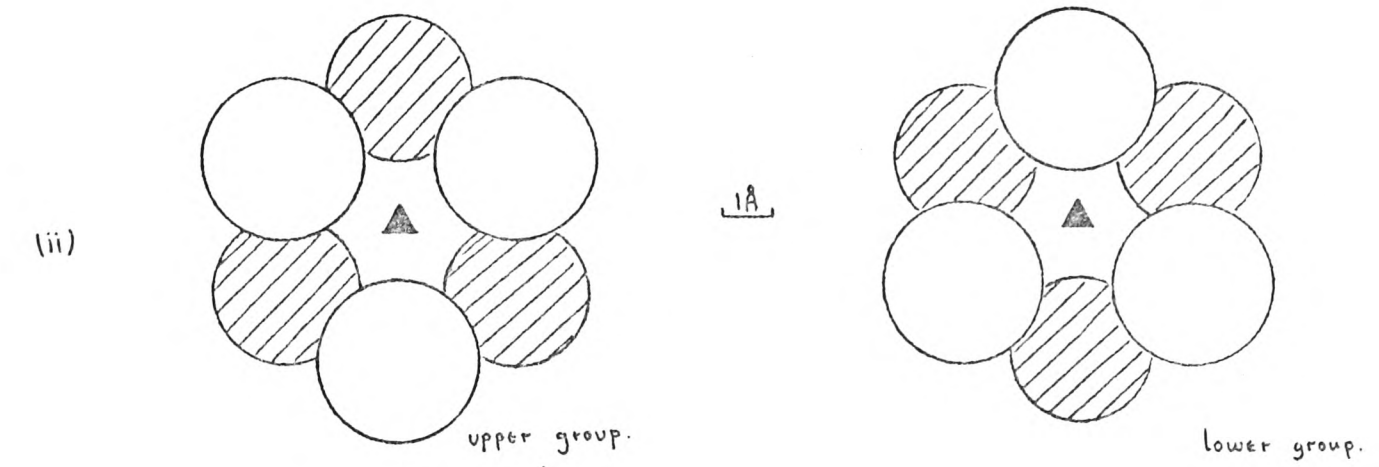
and for the apparently similar arrangement of inter cationic vectors and inter sulphate vectors. The orientation of the double helices was that of Ia (Fig.40) and the cation arrangement is shown in Fig.43. A model was made of three iota double helices surrounding a threefold rotation axis. Beevers models were used with polystyrene balls to represent the cations, their radii being equivalent to the sum of the radii of a potassium ion and an oxygen atom. The structure appeared to be sterically feasible and the cations were in contact with sulphate groups from two double helices, thus binding them together. The upper group of three cations was rather more favourably placed than the lower group as each cation was in contact with three sulphate groups whereas in the lower group each cation was only in contact with two sulphate groups.

In the case of the divalent cation salts it would be reasonable for the cations to occupy the upper group of cation sites, as each cation would be in contact with two sulphate groups of one double helix and with one sulphate group of another double helix, thus binding the two chains of the double helix together and also binding the double helices together. Regardless of packing considerations this would seem to be the most suitable divalent cation site, situated in a slight hollow between the two sulphate groups. For such an arrangement there would be a strong peak at B in the strontium salt Patterson function (Fig.44) and the peak at A would be weak or absent. The strong peak at B is observed but there is still a significant peak at A. This could mean that the divalent cations are distributed between the monovalent cation sites*. The distance of the strontium ion from the helix axis derived/

---

* Space group $P3_2$ would give cation-cation vectors at $z = \frac{1}{3}$.
See note p.110





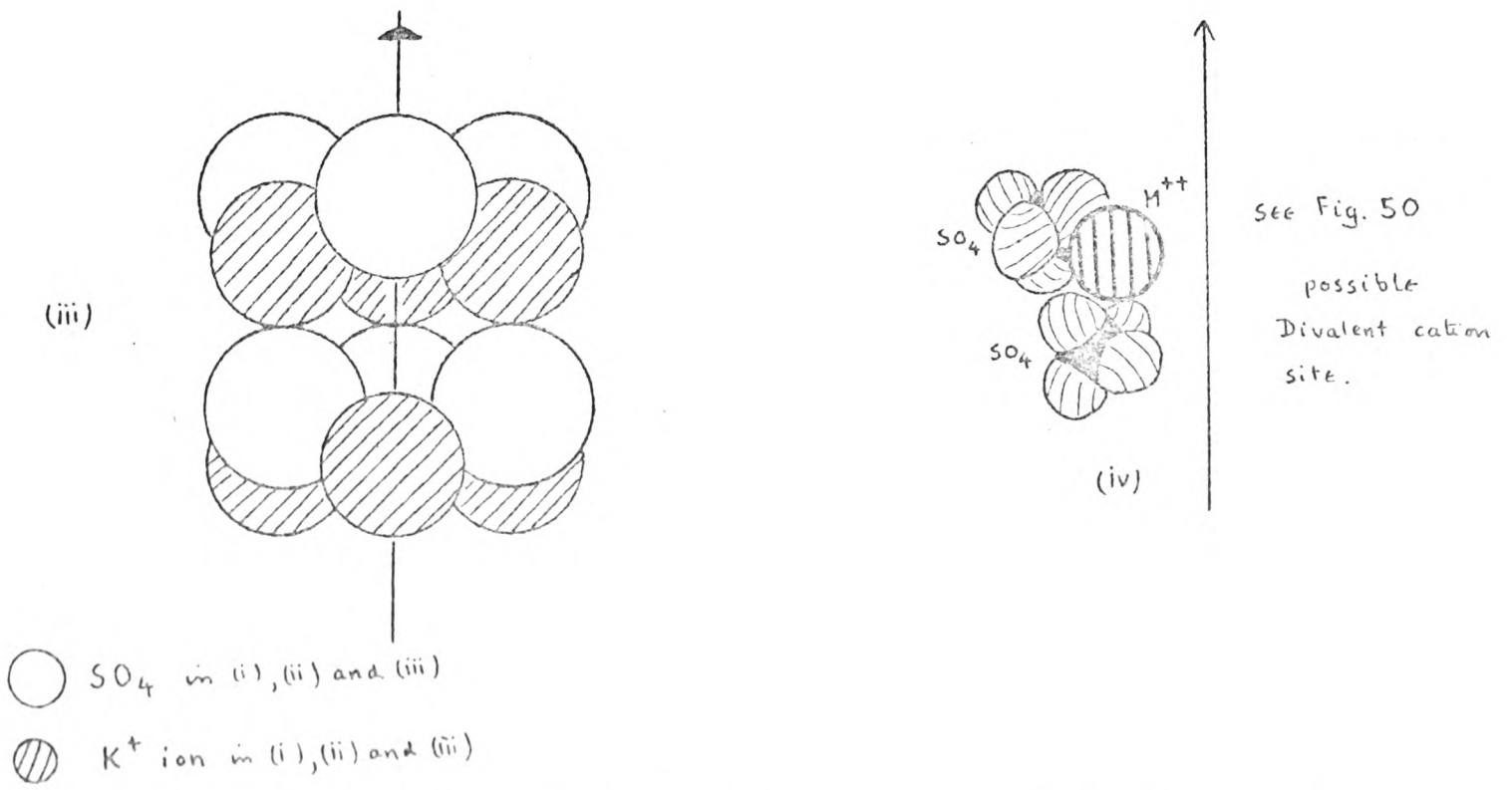


Fig. 43. (i)-(iii) A possible arrangement of sulphate groups and cations round a 3-fold rotation axis. (iv) A possible divalent cation site.

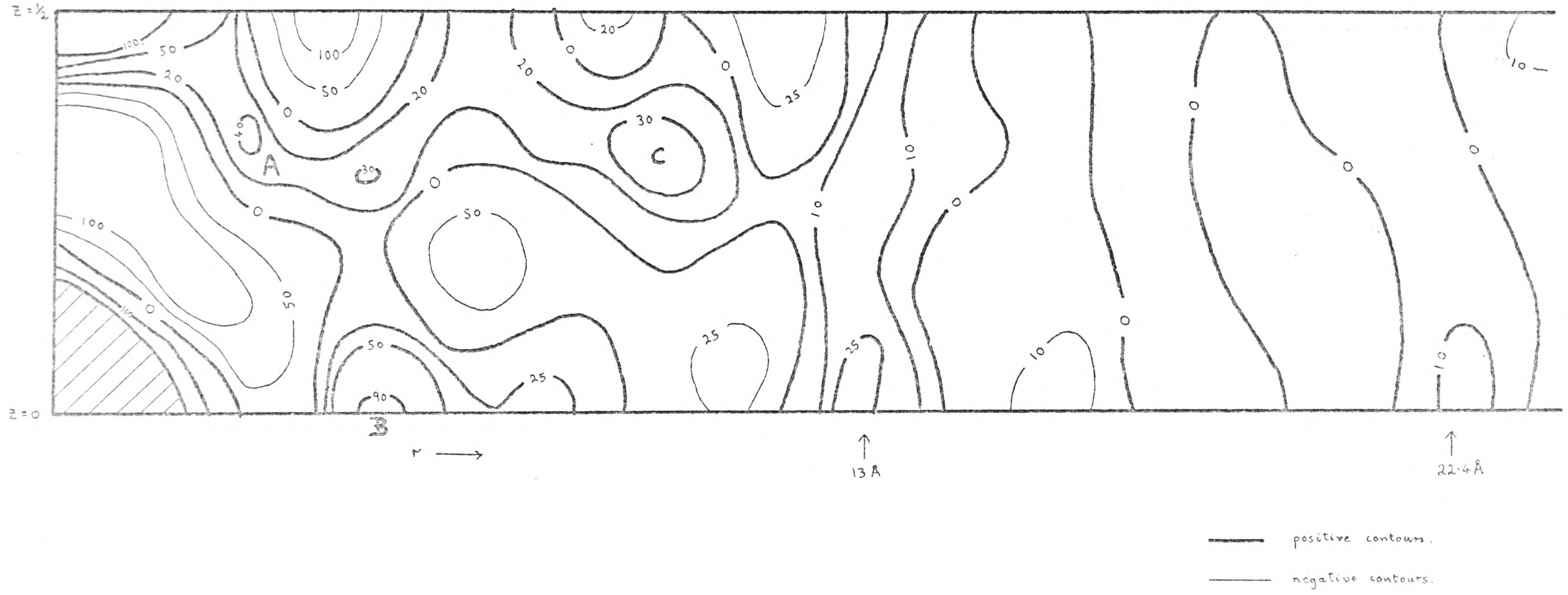


Fig. 44. Cylindrically averaged Patterson function of strontium iota.

derived from the Patterson function, assuming that peak C was the intra double helix cation-cation vector, was rather on the large side to fit into the above structure, being ~ 5.6 Å as compared with ~ 4.8 Å for the potassium salt. The results from the Patterson functions are however rather inconclusive at this stage.

SECTION V: FOURIER TRANSFORM CALCULATIONS AND CONCLUSION

5.1, DIFFRACTION BY HELICAL STRUCTURES

The theory of X-ray diffraction by helical structures has been derived by Cochran, Crick and Vand (75) and by Klug, Crick and Wyckhoff (76).

The transform of a helix of pitch P consisting of repeating units whose projected length onto the helix axis is p is continuous only on planes where

$$\frac{\ell}{\lambda} = \frac{n}{P} + \frac{m}{p}$$

where n and m are integers.
ℓ = the number of the layer line.
$1/c$ = the spacing of the layer lines (Å^{-1}).

The transform on these planes is given by

$$F(R, \Psi, \ell/c) = \sum_n \sum_j f_j \cdot J_n(2\pi . R . r_j) \exp\left[i \left\{ n(\Psi - \phi_j + \pi/2) + 2\pi \ell \frac{z_j}{c} \right\} \right]$$

where the values of n are governed by the selection rule

$$\frac{n}{P} = \frac{\ell}{c} - \frac{m}{p}$$

$J_n(X)$ is the nth order Bessel function of argument X.
$\Psi, R, \ell/c$ are reciprocal space coordinates.
ϕ_j, r_j, z_j are the semi-polar coordinates of the jth atom.
f_j is the scattering factor of the jth atom.

The expression for the transform may be written in the form

$$F(R, \Psi, \ell/c) = \sum_n (A_n + i.B_n) \exp[i.n.\Psi]$$

where
$$A_n = \sum_j f_j \cdot J_n(2.\pi.R.r_j).\cos\left[n(\pi/2 - \phi_j) + 2.\pi.\ell.\frac{z_j}{c}\right]$$

and
$$B_n = \sum_j f_j \cdot J_n(2.\pi.R.r_j).\sin\left[n(\pi/2 - \phi_j) + 2.\pi.\ell.\frac{z_j}{c}\right]$$

Integration over ψ gives the cylindrically averaged transform. The square of the cylindrically averaged transform or the the cylindrically averaged intensity distribution is given by

$$I(R, l/c) = \sum_n (A_n^2 + B_n^2)$$

In calculating cylindrically averaged intensity distributions An and Bn were multiplied by $\exp[-(B\sin^2\theta/\lambda^2)]$ to take into account thermal vibrations of the atoms. B was the overall temperature factor.

A computer program was written by Dr. M.M. Harding to calculate cylindrically averaged intensity distributions. Bessel functions up to the ninth order with a maximum argument of 17.5 were used. Values of the intensity were calculated at intervals of 0.02 in ξ ($= R.\lambda$).

5.2, CALCULATED INTENSITY DISTRIBUTIONS

Cylindrically averaged intensity distributions were calculated for iota model I and for model I with a magnesium ion sited between the two sulphate groups as shown in Fig.43. These are shown in Figs.45 and 46 together with the observed intensity traces for the lithium and magnesium salts. The relative scale factors were estimated by eye. The coordinates of the atoms of the sulphate groups were measured from Beevers models with the sulphate groups placed in sterically reasonable positions. The sulphate groups, especially that on the 3,6 anhydrogalactose, appeared to have some degree of freedom. The calculated transform is rather sensitive to the positions of the sulphate groups as/

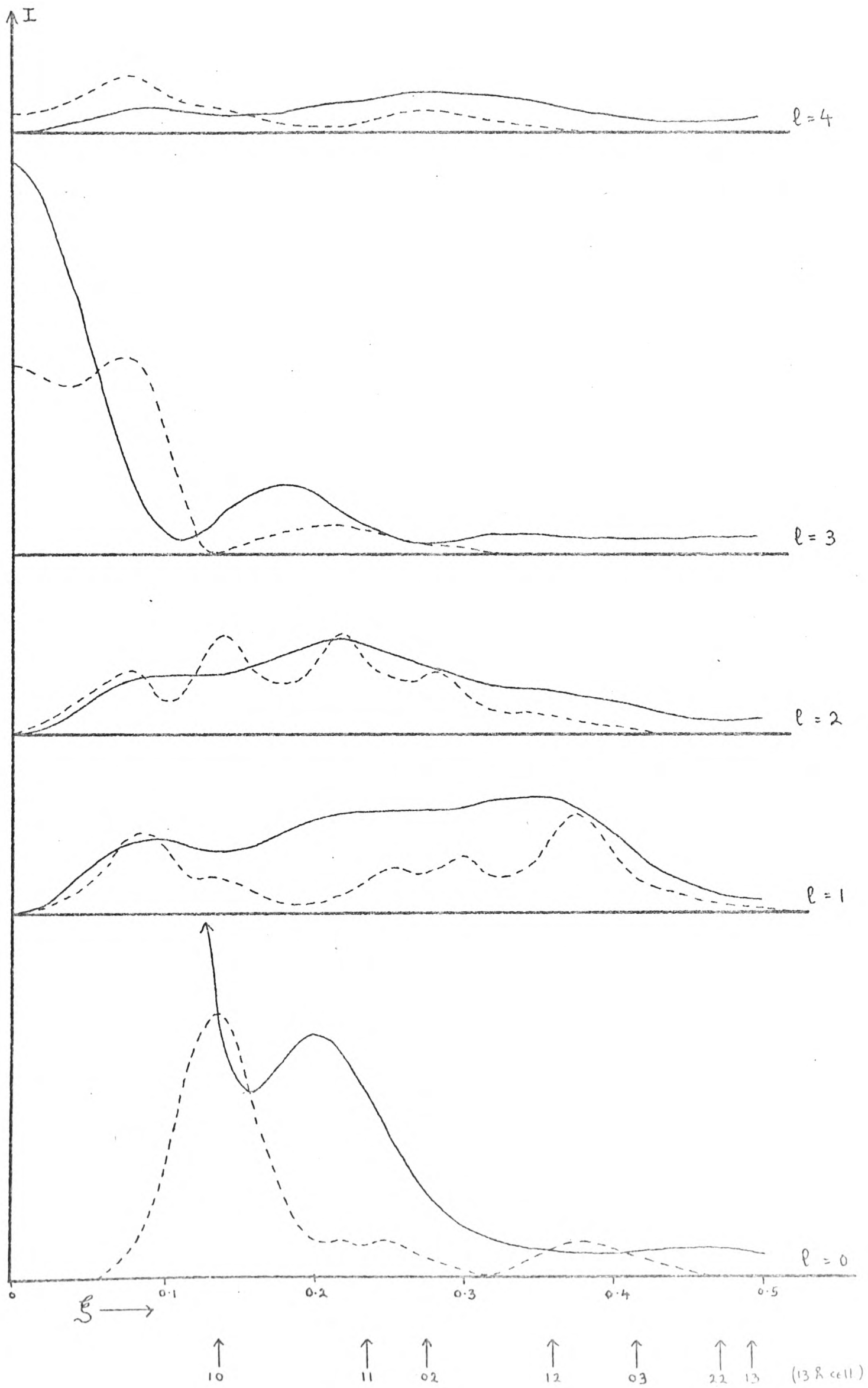


Fig. 45. Calculated (——) and observed (---) intensity distributions for lithium iota.

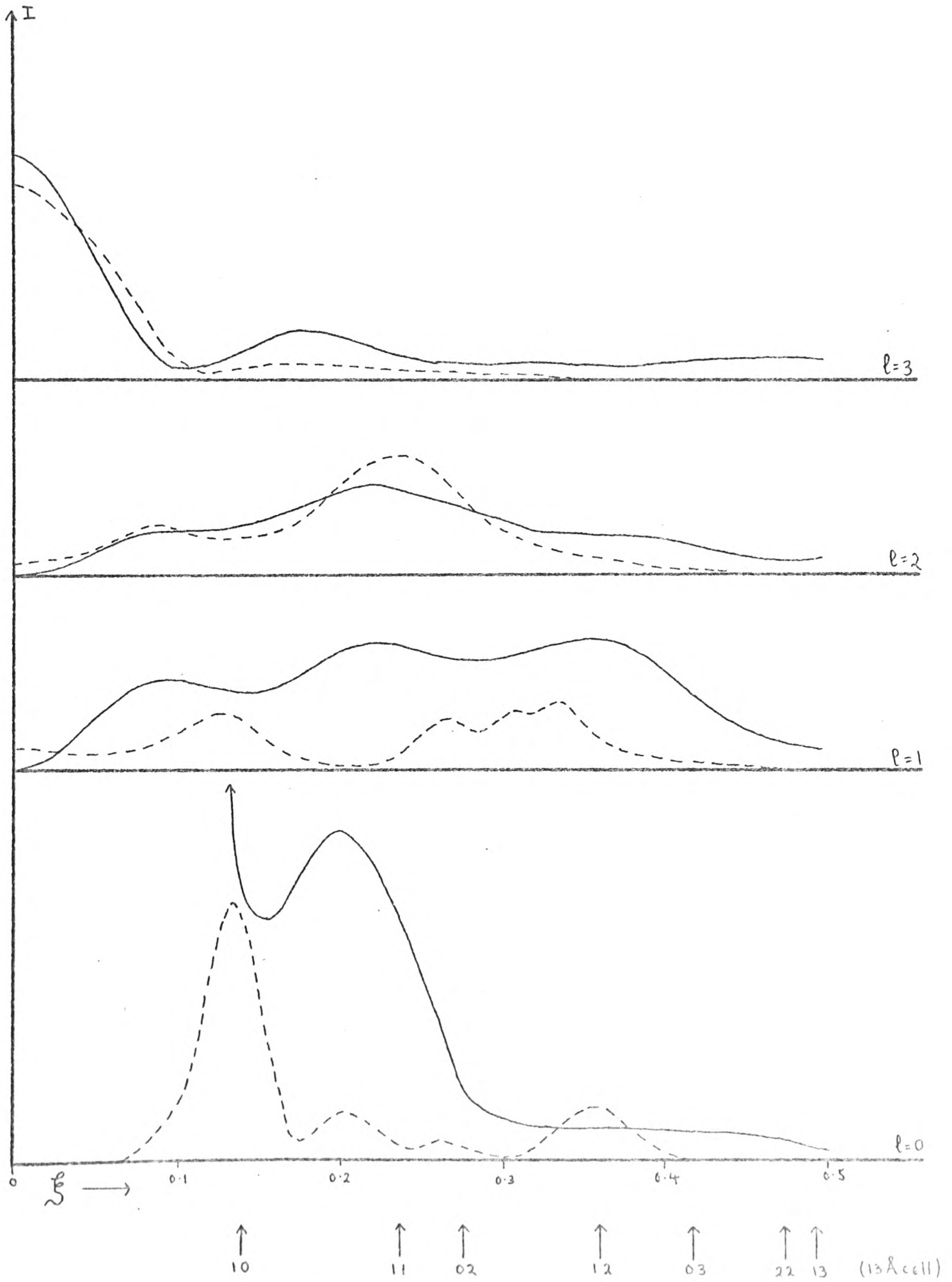


Fig. 46. Calculated (——) and observed (- - -) intensity distributions for magnesium iota.

as they comprise a large part of the scattering matter of the double helix. As an example, the height of the meridional maximum on the third layer line was changed by a factor of 1.35 by changing the z coordinate of one of the sulphate groups by 0.5 Å. The general fit between the observed and calculated intensity traces, allowing for lattice sampling, would appear to be quite reasonable. It should be noted that the cylindrically averaged intensity calculation used does not strictly apply in the present situation where the fibre is at least semi-crystalline and it would be more correct to multiply the molecular transform by the lattice transform before taking the cylindrical average (44). Unfortunately it is not easy to calculate the lattice transform for a semi-crystalline fibre and this has not been attempted yet.

The presence of an intensity maximum at a non reciprocal lattice point on the zero layer at $\xi \sim 0.2$ can be seen to occur where the calculated intensity has a large magnitude.

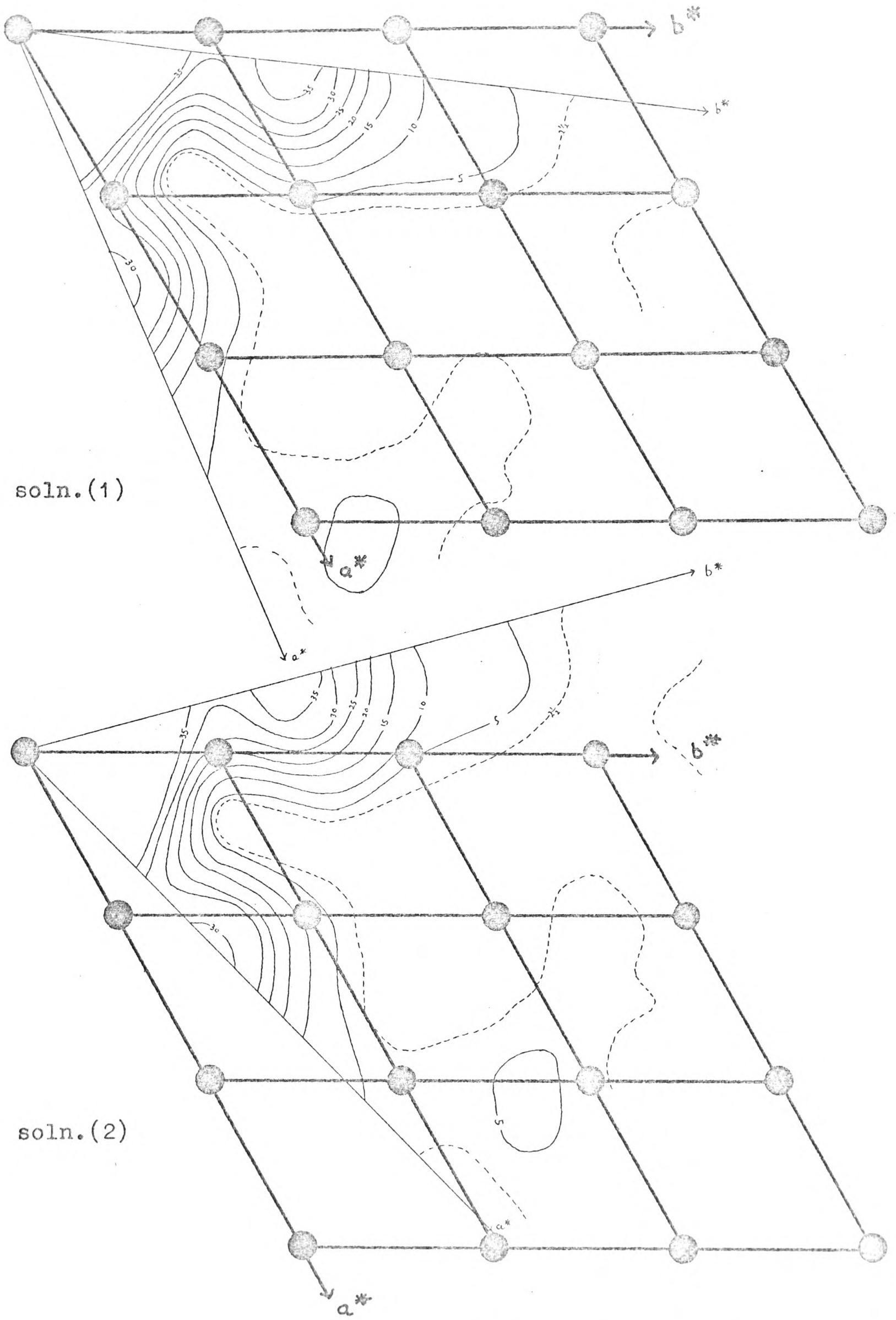
Cylindrically averaged intensity distributions of several other models for iota derived by Dr. D.A. Rees' modelbuild II program were also calculated. These included both left handed and right handed single helices of helix pitch 13.0 Å. The best of these, model H (52), as regards general agreement between calculated and observed intensities, was very similar to model I. The other right handed double helices were all variations on this basic conformation. No left handed double helix solutions were found by either modelbuild I or modelbuild II. The single helix models gave large calculated intensities on the first layer line which/

which were not consistent with the observed diffraction patterns of the lithium and magnesium salts (52).

5.3, ORIENTATION OF THE DOUBLE HELICES AND THE FOURIER TRANSFORM

The X-ray diffraction pattern of a crystal can be considered as the sampling of the Fourier transform of the contents of the unit cell at the reciprocal lattice points. For iota, even in the absence of distortion, this Fourier transform could not be calculated unless the orientations of the double helices passing through the unit cell were known. In the 0001 projection, however, there is a reduced site of unit cell as described above (section 1.7) with only one double helix passing through it and thus it should be possible to find the orientation of the double helix by rotating the Fourier transform of the double helix, projected along the helix axis, over the reciprocal lattice and to find the best fit with the observed intensities. Any solution found will have a two fold ambiguity when related to the full three dimensional structure. The square of the Fourier transform of model I was calculated, for l = o, with respect to hexagonal axes, with a = 89 Å, out to ξ = 0.5. The two dimensional indices used in the following discussion refer to indexing based on the projected cell of side 13 Å. For both the lithium and magnesium salts the 10 reflection is very strong and the 11 and 20 reflections are weak. There is an appreciable intensity at around ξ = 0.36 corresponding to the 30, 12 and 21 reflections. The two most promising orientations are shown in Fig.47 along with the positions of the Fourier transform with respect to the reciprocal lattice.

Though/



soln.(1)

soln.(2)

Fig. 47. Two possible orientations of the Fourier transform with respect to the reciprocal lattice.

Though in both cases the 11 and 20 reflections are rather stronger than desirable, they both occur on sloping parts of the transform close to the low magnitude area. It is possible therefore that small changes in atomic parameters could allow both reciprocal lattice points to fall on low areas of the transform. No orientation would give particularly large intensities for the 30, 12 and 21 reflections but the first solution would be preferable to the second in this respect. The arrangements of the sulphate groups for both orientations are shown in Fig.48. The twofold ambiguity is in whether the axis marked * is a threefold rotation axis or a threefold screw axis. Solution I gives an orientation approximately half way between those described as a and b in the section on the cylindrically averaged Patterson functions (4.4 and Fig.30) and Solution II gives an orientation close to c. Of these orientations Ia and IIc were sterically feasible (see section 4.4).

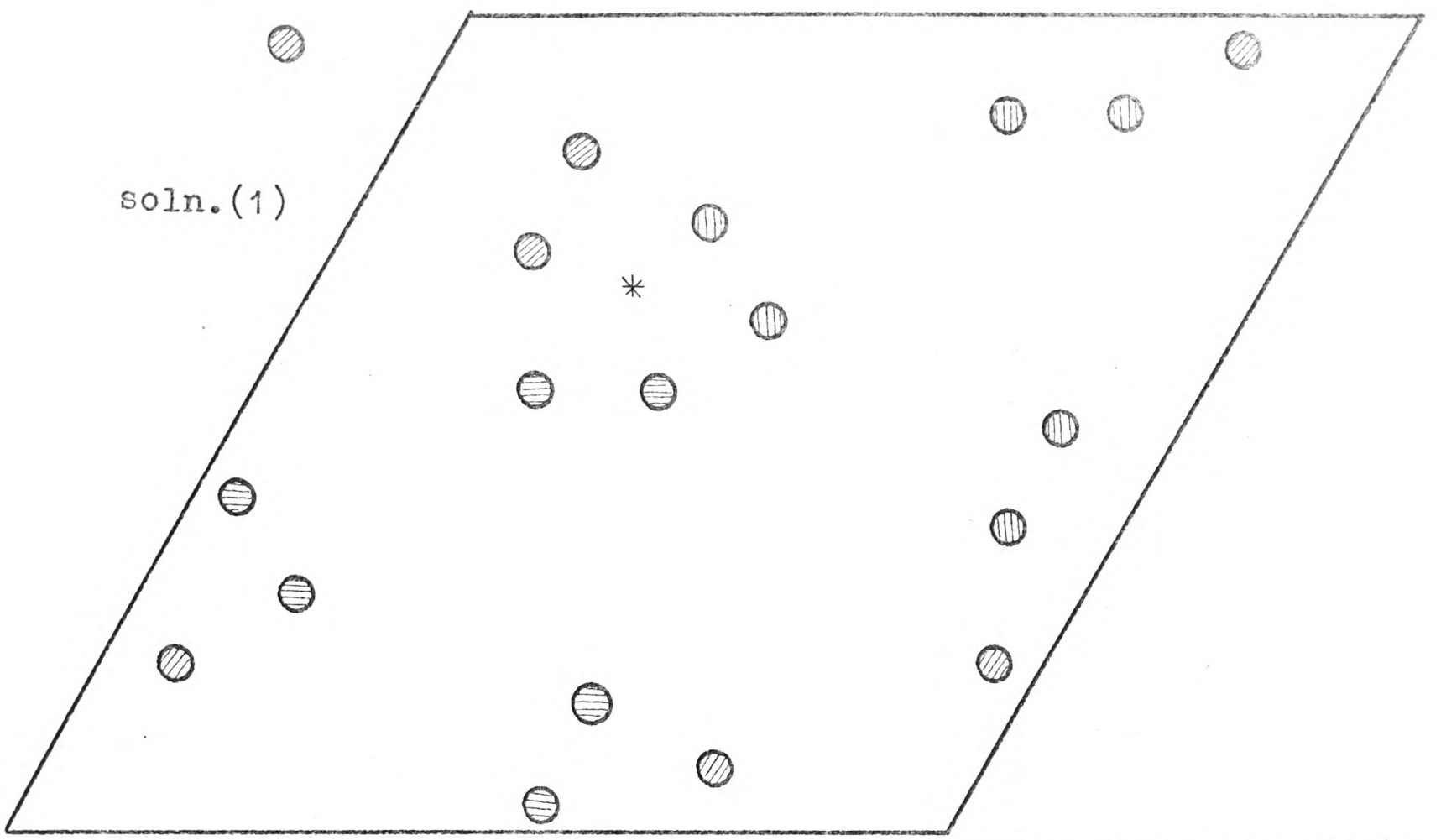
5.4, CONCLUSION

The evidence in favour of the right handed double helix model for iota would appear to the strong. Briefly this evidence may be summarised as follows:

(i) The general features of the X-ray diffraction pattern and the relation between the kappa and iota diffraction patterns are consistent with double helix models.

(ii) The presence of a threefold screw axis is indicated by the presence of meridional reflections only on layers where $l = 3n$.

(iii)/



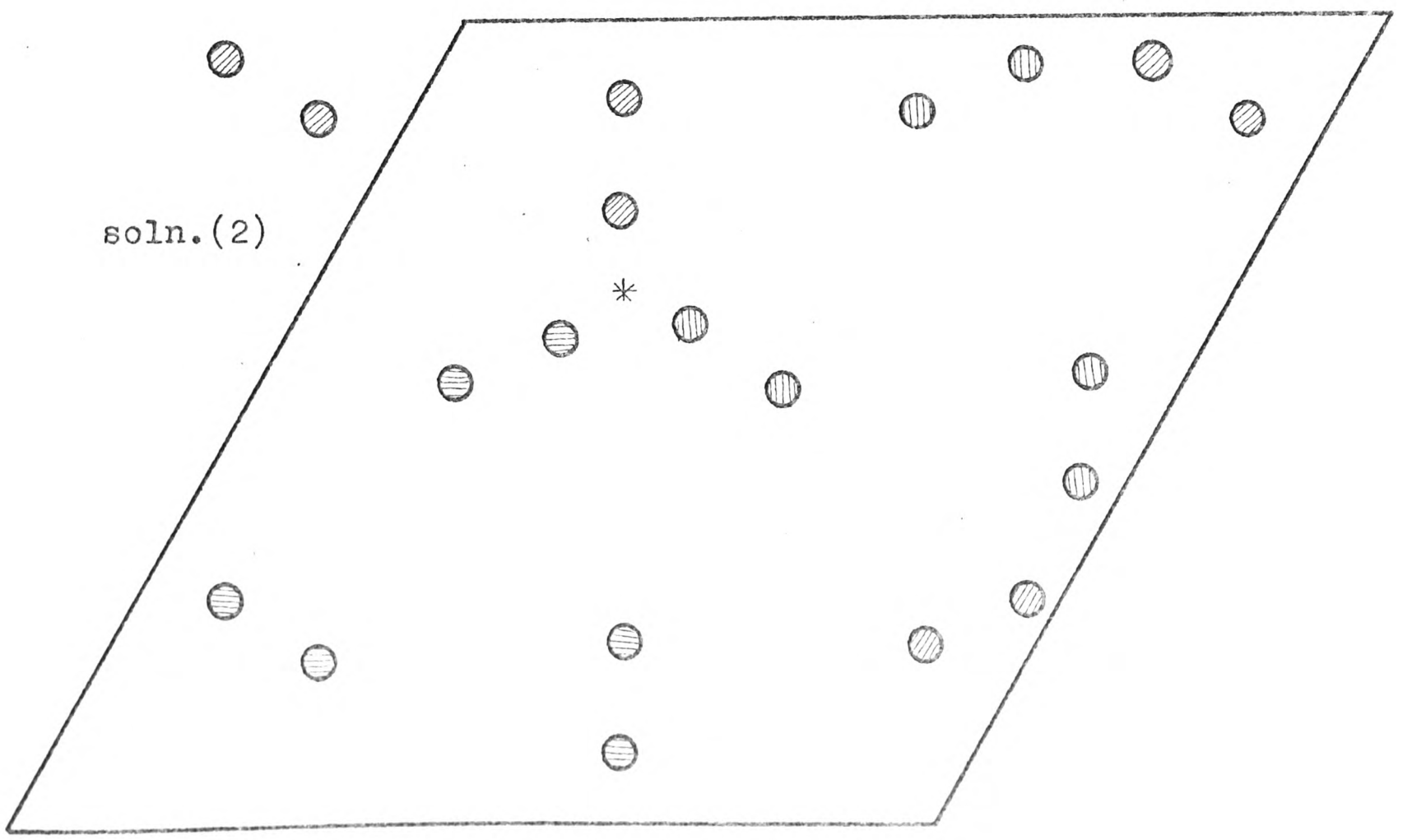


Fig. 48. The arrangements of the sulphate groups in the two possible orientations derived from the Fourier transform.

(iii) The size of the unit cell is consistent with the diameter of the proposed double helices whereas single helices would be difficult to fit in.

(iv) Model-building calculations show that a right handed double helix is sterically feasible.

(v) The presence of a hydrogen bond in the proposed conformation has been confirmed by polarised infra-red radiation studies.

(vi) When a physical model was made, there was ample space to fit on the sulphate groups though they had not been included in the calculation of the chain conformation.

(vii) There is reasonable agreement between the calculated intensity distributions for the proposed model and the observed intensity distributions for the lithium and magnesium salts.

It should be noted that the results of the modelbuilding calculation do not explain why the double helix is formed or why it has threefold screw symmetry and the observed repeat distance. The presence of the hydrogen bond is probably one of the main factors controlling the formation and conformation of the double helix. Crystal packing and the effect of the cations on this is probably also important. This would tend to make a helix which had near threefold screw symmetry into one which had exact threefold screw symmetry as the threefold screw axis is a crystallographic symmetry element.

5.5. SOME SUGGESTED FURTHER WORK

The/

The location of the divalent cations should be an easier problem to solve than that of the monovalent cations, assuming that there is a single divalent cation site rather than the divalent cations being statistically distributed between the monovalent cation sites. It would probably be worth while to measure integrated intensities for the equatorial reflections. The effect on the Fourier transform of changes in sulphate positions should be investigated. If the orientation of the double helices can be determined with reasonable certainty, it might be possible to locate the positions of the cations in projection by calculating a Fourier synthesis projection using phases derived from the double helix with $|F|$ values measured for the heavy atom salts. Unfortunately only a few equatorial reflections can be observed and only those of types h0 and hh (projected cell indices) are single reflections. The hk and kh reflections occur at the same value of ζ but are non-equivalent. Alternatively it might be useful to calculate a Fourier transform with a strontium ion placed in a sterically likely cation site such as that shown in Fig.43. It can be seen from the magnesium and strontium salt photographs (Fig.49) that in the magnesium salt photograph the 10 reflection is very strong and the 11 and 02 reflections are relatively weak, whereas in the strontium salt photograph the 10 reflection is relatively weak and the 11 and 02 reflections are both fairly strong. Even if observed and calculated intensities cannot be accurately matched, it should be possible to account for the general trends in their relative magnitudes.

REFERENCES

1) W.N.Haworth, J.Jackson and F.Smith, J.Chem.Soc.(1940),620

2) R.K.Mauermeyer, E.M.Livingston and R.Michaels, Mikrochimica Acta(1956),11,1600

3) "International Tables for X-ray Crystallography," published for 'The International Union of Crystallography' by The Kynoch Press, Birmingham, England(1952), Vols.I-III.

4) W.Parrish, Acta Cryst.(1960),13,838

5) D.C.Phillips, Acta Cryst.(1954),7,746 and 9,819

6) A.L.Patterson, Z.Kristallogr.(1935),90,517

7) H.Lipson and C.A.Beevers, Proc.Phys.Soc.(1936),48,772

8) C.A.Beevers and H.W.Ehrlich, Z.Kristallogr.(1959),112,414

9) H.Lipson and W.Cochran,"The Determination of Crystal Structures", 'The Crystalline State', Vol.III, ed. Bragg, G.Bell and Sons Ltd.(1966)

10) D.Harker, J.Chem.Phys.(1936),4,381

11) J.Karle and I.L.Karle, Acta Cryst.(1966),21,849

12) I.L.Karle and J.Karle, Acta Cryst.(1964),17,835

13) A.J.C.Wilson, Nature(1942),150,152

14) "Computing Methods in Crystallography", ed. J.S.Rollett, The Pergamon Press(1965)

15) I.L.Karle, K.S.Dragonette and S.A.Brenner, Acta Cryst.(1965),19,713

16) H.Hauptman and J.Karle, Acta Cryst.(1956),9,45

17) J.Karle and H.Hauptman, Acta Cryst.(1956),9,635

18) J.Karle, Acta Cryst.(1968),B24,182

19) L.E.Alexander and G.S.Smith, Acta Cryst.(1962),15,983

20) A.C.T.North, D.C.Phillips and F.S.Mathews, Acta Cryst.(1968),A24,351

21) G.H.Stout and L.H.Jensen, "X-ray Structure Determination", The Macmillan Company, New York(1968)

22) W.R.Busing, K.O.Martin and H.A.Levy(1962), ORFLS., Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

23) D.W.J.Cruickshank, Acta Cryst.(1965),19,153

24) F.R.Ahmed, and D.W.J.Cruickshank, Acta Cryst.(1953), 6,385

25) G.A.Jeffrey and R.D.Rosenstein, Adv.Carb.Chem., (1964),19,7

26) S.S.C.Chu and G.A.Jeffrey, Acta Cryst.(1967),23,1038

27) H.M.Berman, S.S.C.Chu and G.A.Jeffrey, Science(1967),157,1576

28) M.Sundaralingam, J.Amer.Chem.Soc.(1965),87,599

29) J.H.Robertson and B.Sheldrick, Acta Cryst.(1965),19,820

30) C.J.Brown, Sir Gordon Cox and F.S.Llewellyn, J.Chem.Soc.(1966),922

31) G.M.Brown and H.Levy, Science(1963),141,921

32) G.N.Ramachandran, C.Ramakrishnan and V.Sasisekharan, in "Aspects of Protein Structure", ed. G.N.Ramachandran, Academic Press, New York(1963)

33) A.J.de Hoog, H.R.Buys, C.Altona and E.Havinga, Tetrahedron(1969),25,3365

34) A.McL.Mathieson and B.J.Poppleton, Acta Cryst., (1966),21,72

35) C.E.Nordman and K.Nakatsu, J.Amer.Chem.Soc.(1963),85,353

36) M.J.Buerger, "Vector Space", John Wiley and Sons, Inc., New York(1959)

37) D.F.Grant, R.G.Howells and D.Rogers, Acta Cryst., (1957),10,489

38) R.H.Marchessault and A.Sarko, Adv.Carb.Chem.(1967), 22,421

39) J.Blackwell, Biopolymers(1969),7,281

40) A.Sarko and R.H.Marchessault, Science(1966),154,1658

41) S.Arnott and A.J.Wonacott, J.Mol.Biol.(1966),21,371

42) S.Arnott and S.D.Dover, J.Mol.Biol.(1967),30,209

43) S.Arnott, S.D.Dover and A.Eliott, J.Mol.Biol., (1967),30,201

44) K.C.Holmes and D.M.Blow, Methods Biochem. Analysis, (1965),13,113; (Interscience reprint, John Wiley and Sons, Inc.(1966))

45) D.A.Rees, "The Shapes of Molecules", Oliver and Boyd Ltd., Edinburgh(1967)

46) D.A.Rees, Ann.Reports(1965),62,469

47) N.S.Anderson, T.C.S.Dolan and D.A.Rees, Nature(1965),205,1060

48) N.S.Anderson, T.C.S.Dolan and D.A.Rees, J.Chem.Soc., (1968),C,596

49) N.S.Anderson, T.C.S.Dolan and D.A.Rees, unpublished.

50) D.A.Rees, Adv.Carb.Chem.(1969),24,267

51) D.A.Rees, J.Chem.Soc.(1961),5168

52) N.S.Anderson, J.W.Campbell, M.M.Harding, D.A.Rees and J.W.B.Samuel, J.Mol.Biol.(1969),45,85

53) W.Fuller, F.Hutchison, M.Spencer and M.H.F.Wilkins, J.Mol.Biol.(1967),27,507

54) S.T.Bayley, Biochim.Biophys.Acta(1955),17,194

55) E.D.T.Atkins, K.D.Parker and R.D.Preston, Proc.Roy.Soc.(1969),B173,209

56) T.J.Painter, Canad.J.Chem(1960),38,112

57) N.S.Anderson, T.C.S.Dolan, A.Penman, D.A.Rees, G.P.Mueller, D.J.Stancioff and N.F.Stanley, J.Chem.Soc.(1968),C,602

58) A.A.McKinnon, D.A.Rees and F.B.Williamson, Chem.Comm.(1969),701

59) R.E.Franklin and R.G.Gosling, Acta Cryst.(1953),6,678

60) B.K.Vainstein, "Diffraction of X-rays by Chain Molecules", English edition, Elsevier publishing company, Amsterdam-London-New York(1966)

61) M.Sundaralingam, Biopolymers(1968),6,189

62) D.A.Rees, J.Chem.Soc.(1969),B,217

63) D.W.Jones, J.Polymer Sci.(1958),32,371

64) W.J.Settineri and R.H.Marchessault, J.Polymer Sci., (1965),C11,253

65) A.Sarko and R.H.Marchessault, J.Amer.Chem.Soc., (1967),89,6454

66) C.V.Durell and A.Robson, "Advanced Algebra, Vol.II", G.Bell and Sons Ltd., London(1966)

67) A.I.Kitaygorodsky, Tetrahedron(1961),14,230

68) V.S.R.Rao, P.R.Sundararajan, C.Ramakrishnan and G.N.Ramachandran, in "Conformation of Biopolymers", ed. G.N.Ramachandran, Academic Press, New York(1967),Vol.II

69) D.A.Rees and R.J.Skerrett, Carbohydrate Res.(1968), 7,334

70) C.A.Beevers, J.Chem.Educ.(1965),42,273

71) C.H.MacGillavry and E.M.Bruins, Acta Cryst.(1948),1,156

72) "Handbook of Mathematical Functions", ed. M.Abramowitz and I.A.Stegun, Dover publications Inc., New York(1965)

73) M.J.Buerger, Proc.Nat.Acad.Sci.Wash.(1942),28,281

74) D.M.Blow, Proc.Roy.Soc.(1958),A247,302

75) W.Cochran, F.H.C.Crick and V.Vand, Acta Cryst.(1952),5,581

76) A.Klug, F.H.C.Crick and H.W.Wyckhoff, Acta Cryst., (1958),11,199

APPENDIX I

COMPUTER PROGRAM FOR MODELBUILD I

(IN ATLAS AUTOCODE, EDINBURGH UNIVERSITY ISO VERSION)

```
 %BEGIN

%REAL     CO,SI,ALPHA,PSI,PHI,AA,AAA,BX',BY',BZ',LD,HR,COG,SIG
%REAL     CSRO,SNRO,NN,BAB,ST,CT
%REAL     T1,T2,T3,TTA,TTB,XC3,YC3,ZC3
%REAL     COA,SIA,BB,BBB,COB,SIB,COT,SIT,OMEGA,COW,CA,CB,CC,CD,CE
%REAL     XC4,YC4,ZC4,X1,X2,X3,X4,Y1,Y2,Y3,Y4,Z1,Z2,Z3,Z4
%REAL     AX',AY',AZ',S1,S2,RR,RRR
%INTEGER  AN3,BN3,HAND
%ARRAY    DHM,DHM1,DHM2,DHM3(1:3,1:3)
%REAL     ABA,SID,COD
%REAL     CA0,CA1,CA2,CA3,CA4,II,JJ,GG,HH,DTA1,DTA2,P3,Q3
%REAL     P,Q,LAM,P1,Q1,B1,C1,B2,C2,DT1,DT2,CSR1,CSR2,SNR1,SNR2
%REAL     RR1,RR2,DD,ANG,VANG,ALPHA',P',Q'
%INTEGER  I,I',PSI1,DPSI,PSI2,L,NA,AN1,AN2,I2
%INTEGER  J,J',PHI1,DPHI,PHI2,K,NB,BN1,BN2
%ARRAY    M1,M2,M3,M4,M5,M21,M321,MM1,MM2,MM3,AM1,AM2,AMM(1:3,1:3)
%ARRAY    ABX1,ABY1,ABZ1,BX1,BY1,BZ1(1:25)
%ARRAY    ABX2,ABY2,ABZ2,BX2,BY2,BZ2(1:25)
%ARRAY    ABX3,ABY3,ABZ3,BX3,BY3,BZ3(1:25)
%ARRAY    ABX4,ABY4,ABZ4,BX4,BY4,BZ4(1:25)
%ARRAY    AX1,AY1,AZ1(1:25)

%ROUTINESPEC    MATRIXPROD(%REALARRAYNAME C,A,B)
%REALFNSPEC     QUADRANT(%REAL SINANG,COSANG)

%ARRAY VDW,VDW'(1:6,1:6)
%INTEGER NAT,AN4,BN4,AN5,BN5
%INTEGER %ARRAY AT,AT'(1:25)

%CAPTION ~~~~POLYSACCHARIDE␣MODELBUILDING␣PROGRAM␣FOR␣REPEATING
%CAPTION ␣POLYSACCHARIDES␣OF␣TYPE␣(-A-B-)N
%CAPTION ~~~VERSION␣FOR␣RIGHT␣HANDED␣OR␣LEFT␣HANDED␣HELICES␣OF
%CAPTION ␣ANY␣SCREW␣SYMMETRY␣AND␣HELIX␣REPEAT
%CAPTION ~~~CONTACTS␣AT␣BOTH␣LINKAGES␣ARE␣CHECKED
%CAPTION ~~~(J.W.CAMPBELL␣29/5/69␣VERSION)
1:READ(I)
->2 %IF I=-1
READ(AX1(I),AY1(I),AZ1(I),AT(I))
NA=I
->1
2:READ(AN1,AN2)
12:READ(I)
->3 %IF I=-1
READ(BX1(I),BY1(I),BZ1(I),AT'(I))
NB=I
->12
3:READ(BN1,BN2)
READ(ALPHA')
ALPHA=(π-(π*ALPHA')/(180))
READ(PHI1,DPHI,PHI2)
READ(PSI1,DPSI,PSI2)
READ(HR)
READ(NN)
READ(OMEGA)
READ(AN3,BN3,AN4,BN4,AN5,BN5)
```

```
READ(NAT)
%CAPTION ~~VAN␣DER␣WAALS␣CONTACT␣DISTANCES
%CYCLE I=1,1,NAT
NEWLINE
%CYCLE J=1,1,NAT
READ(VDW'(I,J))
PRINT(VDW'(I,J),3,2)
VDW(I,J)=(VDW'(I,J))↑2
%REPEAT
%REPEAT
HAND=1 %IF NN>=0
HAND=-1 %IF NN<0
NN=!NN!
PHI1=HAND*PHI1;DPHI=HAND*DPHI;PHI2=HAND*PHI2
PSI1=HAND*PSI1;DPSI=HAND*DPSI;PSI2=HAND*PSI2
%CYCLE I=1,1,NA
AY1(I)=HAND*AY1(I)
%REPEAT
%CYCLE I=1,1,NB
BY1(I)=HAND*BY1(I)
%REPEAT
->7701 %IF HAND=-1
%CAPTION ~~~POLYSACCHARIDE␣CHAIN␣RIGHT␣HANDED␣HELIX
->7702
7701:%CAPTION ~~~POLYSACCHARIDE␣CHAIN␣LEFT␣HANDED␣HELIX
7702:PRINT(NN,2,3)
%CAPTION ␣RESIDUES␣PER␣HELIX␣REPEAT
%CAPTION ~~~HELIX␣REPEAT␣=
PRINT(HR,2,3)
%CAPTION ␣ANGSTROMS
CO=COS(ALPHA)
SI=SIN(ALPHA)
M2(1,1)= CO;M2(1,2)= 0;M2(1,3)= SI
M2(2,1)=  0;M2(2,2)= 1;M2(2,3)=  0
M2(3,1)=-SI;M2(3,2)= 0;M2(3,3)= CO

%CYCLE L=PSI1,DPSI,PSI2
PSI=(π*L)/(180)
CO=COS(PSI)
SI=SIN(PSI)
M1(1,1)=  1;M1(1,2)=  0;M1(1,3)=  0
M1(2,1)=  0;M1(2,2)= CO;M1(2,3)=-SI
M1(3,1)=  0;M1(3,2)= SI;M1(3,3)= CO
MATRIXPROD(M21,M2,M1)

%CYCLE K=PHI1,DPHI,PHI2
NEWLINES(2)
%CAPTION ~ROTATION␣PARAMETER␣PHI1␣=;PRINT(K*HAND,3,0)
%CAPTION ~ROTATION␣PARAMETER␣PSI1␣=;PRINT(L*HAND,3,0)
PHI=(π*K)/(180)
CO=COS(PHI)
SI=SIN(PHI)
M3(1,1)=  1;M3(1,2)=  0;M3(1,3)=  0
M3(2,1)=  0;M3(2,2)= CO;M3(2,3)=-SI
M3(3,1)=  0;M3(3,2)= SI;M3(3,3)= CO
MATRIXPROD(M321,M3,M21)
```

```
%CYCLE I=1,1,NA
ABX1(I)=M321(1,1)*AX1(I)+M321(1,2)*AY1(I)+M321(1,3)*AZ1(I)
ABY1(I)=M321(2,1)*AX1(I)+M321(2,2)*AY1(I)+M321(2,3)*AZ1(I)
ABZ1(I)=M321(3,1)*AX1(I)+M321(3,2)*AY1(I)+M321(3,3)*AZ1(I)
%REPEAT
%CYCLE I=1,1,NA
ABX2(I)=ABX1(I)-BX1(BN1)
ABY2(I)=ABY1(I)-BY1(BN1)
ABZ2(I)=ABZ1(I)-BZ1(BN1)
%REPEAT
%CYCLE I=1,1,NB
BX2(I)=BX1(I)-BX1(BN1)
BY2(I)=BY1(I)-BY1(BN1)
BZ2(I)=BZ1(I)-BZ1(BN1)
%REPEAT
AA=SQRT((ABX2(AN1))↑2+(ABY2(AN1))↑2)
CO=ABX2(AN1)/(AA)
SI=ABY2(AN1)/(AA)
MM1(1,1)= CO;MM1(1,2)= SI;MM1(1,3)=  0
MM1(2,1)=-SI;MM1(2,2)= CO;MM1(2,3)=  0
MM1(3,1)=  0;MM1(3,2)=  0;MM1(3,3)=  1
AAA=SQRT((ABX2(AN1))↑2+(ABY2(AN1))↑2+(ABZ2(AN1))↑2)

CO=(AA)/(AAA)
SI=ABZ2(AN1)/(AAA)
MM2(1,1)= CO;MM2(1,2)=  0;MM2(1,3)= SI
MM2(2,1)=  0;MM2(2,2)=  1;MM2(2,3)=  0
MM2(3,1)=-SI;MM2(3,2)=  0;MM2(3,3)= CO
MATRIXPROD(M4,MM2,MM1)
BX'=M4(1,1)*BX2(BN2)+M4(1,2)*BY2(BN2)+M4(1,3)*BZ2(BN2)
BY'=M4(2,1)*BX2(BN2)+M4(2,2)*BY2(BN2)+M4(2,3)*BZ2(BN2)
BZ'=M4(3,1)*BX2(BN2)+M4(3,2)*BY2(BN2)+M4(3,3)*BZ2(BN2)
AA=SQRT((BY')↑2+(BZ')↑2)
CO=(BY')/(AA)
SI=(BZ')/(AA)
MM3(1,1)=  1;MM3(1,2)=  0;MM3(1,3)=  0
MM3(2,1)=  0;MM3(2,2)= CO;MM3(2,3)= SI
MM3(3,1)=  0;MM3(3,2)=-SI;MM3(3,3)= CO
MATRIXPROD(M5,MM3,M4)
%CYCLE I=1,1,NA
ABX3(I)=M5(1,1)*ABX2(I)+M5(1,2)*ABY2(I)+M5(1,3)*ABZ2(I)
ABY3(I)=M5(2,1)*ABX2(I)+M5(2,2)*ABY2(I)+M5(2,3)*ABZ2(I)
ABZ3(I)=M5(3,1)*ABX2(I)+M5(3,2)*ABY2(I)+M5(3,3)*ABZ2(I)
%REPEAT
%CYCLE I=1,1,NB
BX3(I)=M5(1,1)*BX2(I)+M5(1,2)*BY2(I)+M5(1,3)*BZ2(I)
BY3(I)=M5(2,1)*BX2(I)+M5(2,2)*BY2(I)+M5(2,3)*BZ2(I)
BZ3(I)=M5(3,1)*BX2(I)+M5(3,2)*BY2(I)+M5(3,3)*BZ2(I)
%REPEAT
AAA=(ABX3(AN1))↑2+(ABY3(AN1))↑2+(ABZ3(AN1))↑2
LD=SQRT(AAA)
%CAPTION ~LENGTH OF DISACCHARIDE =
PRINT(LD,1,4)
->100 %IF LD<(HR/NN)
```

```
AAA=(BX3(BN2))↑2+(BY3(BN2))↑2
AA=SQRT(AAA)
C□G=(BX3(BN2))/AA
SIG=(BY3(BN2))/AA
AAA=(ABY3(AN2))↑2+(ABZ3(AN2))↑2
AA=SQRT(AAA)

C□=ABY3(AN2)/AA
SI=ABZ3(AN2)/AA

C□T=((4*(HR↑2))/(NN↑2) + 2*(LD↑2 - (HR↑2)/(NN↑2))*(1 + C□S%C
((2*π)/NN)))/(2*(LD↑2)) - 1
SIT=SQRT(1-((C□T)↑2))
BAB=(3*HR/NN)↑2 + ((LD↑2-(HR↑2)/(NN↑2))*((SIN(3*π/NN))↑2))/(%C
(SIN(π/NN))↑2)
ABA=(BAB -(LD↑2)*(1+2*C□T)↑2)/((LD↑2)*(SIT↑2))
C□D=-(ABA-2)/2
SID=SQRT(1-C□D↑2)
C□A=C□*C□D+SI*SID
SIA=SI*C□D-C□*SID
BBB=(ABX3(AN1)-ABX3(AN2))↑2+AAA
BB=SQRT(BBB)
C□B=(ABX3(AN1)-ABX3(AN2))/BB
SIB=AA/BB

AA=(π*□MEGA)/(180)
C□W=C□S(AA)
CA=(SIB*SIG*C□A)-(SIB*SIG*C□T*C□A)
CB=(SIB*SIG*C□T*SIA)-(SIB*SIG*SIA)
CC=(SIB*SIT*C□G*SIA)
CD=(SIG*C□B*SIT)+(SIB*C□G*SIT*C□A)
CE=(C□B*C□G*C□T)+(C□W)-(SIG*SIB*C□A)
CA0=CA↑2 + CB↑2
CA1=((2*CB*CC)-(2*CA*CD))/4
CA2=(CC↑2 - CB↑2 + 2*CE*CA + CD↑2)/6
CA3=(-(2*CB*CC + 2*CE*CD))/4
CA4=CE↑2 - CC↑2

II=(CA0*((CA0*CA4)-(4*CA1*CA3)+(3*(CA2↑2))))/CA0
JJ=((((CA2↑2)-(CA0*CA4))*((CA1↑2)-(CA0*CA2)))-((CA1*CA2)-(CA0*CA3))%C
↑2)/CA0

GG=JJ/4
HH=-II/12
DTA1=II↑3 - 27*(JJ↑2)
DTA2=GG↑2 +  4*(HH↑3)
->175 %UNLESS DTA2>0
P3=(GG+SQRT(DTA2))/2
Q3=(GG-SQRT(DTA2))/2
->60 %UNLESS P3=0
P=0
-> 65
60:P'=(1/3)*(L□G(!P3!))
P= EXP(P') %IF P3 > 0
P=-EXP(P') %IF P3 < 0
65:->70 %UNLESS Q3 = 0
Q=0
-> 75
```

```
70:Q'=(1/3)*(LOG(!Q3!))
Q= EXP(Q') %IF Q3 > 0
Q=-EXP(Q') %IF Q3 < 0
75:LAM=-P-Q
P1=SQRT((CA0*LAM)+CA1↑2-(CA0*CA2))
Q1=((2*CA1*LAM)+(CA1*CA2)-(CA0*CA3))/P1
B1=2*(CA1-P1)
C1=(CA2+(2*LAM)-Q1)
B2=2*(CA1+P1)
C2=(CA2+(2*LAM)+Q1)
DT1=(B1↑2-(4*CA0*C1))
DT2=(B2↑2-(4*CA0*C2))
->20 %IF DT1>=0 %AND DT2>=0
->25 %IF DT1 <0 %AND DT2 <0
->30 %IF DT1>=0 %AND DT2 <0
->35 %IF DT2>=0 %AND DT1 <0
30:CSR1=(-B1+SQRT(DT1))/(2*CA0)
   CSR2=(-B1-SQRT(DT1))/(2*CA0)
->40
35:CSR1=(-B2+SQRT(DT2))/(2*CA0)
   CSR2=(-B2-SQRT(DT2))/(2*CA0)
->40
20:%CAPTION ~DISCRIMINANTS␣DT1␣AND␣DT2␣BOTH␣POSITIVE
->200
25:%CAPTION ~DISCRIMINANTS␣DT1␣AND␣DT2␣BOTH␣NEGATIVE
->200
40:SNR1=((CD*CSR1)-(CA*((CSR1)↑2))-CE)/((CB*CSR1)+CC)
   SNR2=((CD*CSR2)-(CA*((CSR2)↑2))-CE)/((CB*CSR2)+CC)
DD=SQRT(LD↑2 - (HR↑2)/(NN↑2))
AA=2*(LD↑2-(HR↑2)/(NN↑2))*(1+COS(2*π/NN))
AAA=SQRT(4*((HR↑2)/(NN↑2)) + AA)
SNR0=(2*HR)/(NN*AAA)
CSR0=SQRT(1-SNR0↑2)
CT=SQRT((1 + C0T)/2)
ST=SQRT(1 - CT↑2)

AM2(1,1)=-ST
AM2(1,2)=-CT
AM2(1,3)=0

AM2(2,1)= CT*CSR0
AM2(2,2)=-ST*CSR0
AM2(2,3)=-SNR0

AM2(3,1)= CT*SNR0
AM2(3,2)=-ST*SNR0
AM2(3,3)= CSR0
```

```
%CYCLE I2=1,1,2
NEWLINE
->505 %IF I2=2
%CAPTION ~FIRST␣SOLUTION
NEWLINE
->510
505:%CAPTION ~SECOND␣SOLUTION
NEWLINE
510:CO=CSR1 %IF I2=1
    SI=SNR1 %IF I2=1
    CO=CSR2 %IF I2=2
    SI=SNR2 %IF I2=2

AM1(1,1)=  1;AM1(1,2)=  0;AM1(1,3)=  0
AM1(2,1)=  0;AM1(2,2)= CO;AM1(2,3)=-SI
AM1(3,1)=  0;AM1(3,2)= SI;AM1(3,3)= CO
MATRIXPROD(AMM,AM2,AM1)
%CAPTION ~AT.NO␣␣␣␣X␣␣␣␣␣␣␣Y␣␣␣␣␣␣␣Z␣␣␣␣␣␣␣R␣␣␣␣␣␣␣PHI
NEWLINE
AA=DD/(2*(SIN(π/NN)))
%CYCLE I=1,1,NA
NEWLINE
PRINT(I,2,0)
ABX4(I)=AMM(1,1)*ABX3(I)+AMM(1,2)*ABY3(I)+AMM(1,3)*ABZ3(I) + AA
ABY4(I)=AMM(2,1)*ABX3(I)+AMM(2,2)*ABY3(I)+AMM(2,3)*ABZ3(I)
ABZ4(I)=AMM(3,1)*ABX3(I)+AMM(3,2)*ABY3(I)+AMM(3,3)*ABZ3(I)
PRINT(ABX4(I),.6,3)
PRINT(HAND*ABY4(I),3,3)
PRINT(ABZ4(I),3,3)
BB=SQRT((ABX4(I))↑2 + (ABY4(I))↑2)
PRINT(BB,3,3)
BBB=QUADRANT(((HAND*ABY4(I))/BB),((ABX4(I))/BB))
PRINT(((180*BBB)/π),3,3)
%REPEAT
NEWLINE
%CYCLE I=1,1,NB
NEWLINE
PRINT(I,2,0)
BX4(I)=AMM(1,1)*BX3(I)+AMM(1,2)*BY3(I)+AMM(1,3)*BZ3(I) + AA
BY4(I)=AMM(2,1)*BX3(I)+AMM(2,2)*BY3(I)+AMM(2,3)*BZ3(I)
BZ4(I)=AMM(3,1)*BX3(I)+AMM(3,2)*BY3(I)+AMM(3,3)*BZ3(I)
PRINT(BX4(I),6,3)
PRINT(HAND*BY4(I),3,3)
PRINT(BZ4(I),3,3)
BB=SQRT((BX4(I))↑2 + (BY4(I))↑2)
PRINT(BB,3,3)
BBB=QUADRANT(((HAND*BY4(I))/BB),((BX4(I))/BB))
PRINT(((180*BBB)/π),3,3)
%REPEAT

XC3=BX4(BN2)*(COS(2*π/NN)) - BY4(BN2)*(SIN(2*π/NN))
YC3=BY4(BN2)*(COS(2*π/NN)) + BX4(BN2)*(SIN(2*π/NN))
ZC3=BZ4(BN2) + HR/NN
```

```
T1=(ABX4(AN1)-ABX4(AN2))↑2 + (ABY4(AN1)-ABY4(AN2))↑2 + (ABZ4(AN1)-%C
ABZ4(AN2))↑2
T2=(ABX4(AN1)-XC3)↑2 + (ABY4(AN1)-YC3)↑2 + (ABZ4(AN1)-ZC3)↑2
T3=(ABX4(AN2)-XC3)↑2 + (ABY4(AN2)-YC3)↑2 + (ABZ4(AN2)-ZC3)↑2
TTA=(T1+T2-T3)/(2*(SQRT(T1))*(SQRT(T2)))
TTB=ARCCOS(TTA)
%CAPTION ~BRIDGE␣ANGLE␣=
PRINT(((180*TTB)/π),3,2)
XC4=BX4(BN3)*(COS(2*π/NN)) - BY4(BN3)*(SIN(2*π/NN))
YC4=BY4(BN3)*(COS(2*π/NN)) + BX4(BN3)*(SIN(2*π/NN))
ZC4=BZ4(BN3) + HR/NN


%CYCLE J'=1,1,2
X1=0;X2=0;X3=0
->300 %IF J'=2
X2=ABX4(AN2)-ABX4(AN1)
Y2=ABY4(AN2)-ABY4(AN1)
Z2=ABZ4(AN2)-ABZ4(AN1)
X3=XC3-ABX4(AN1)
Y3=YC3-ABY4(AN1)
Z3=ZC3-ABZ4(AN1)
X4=ABX4(AN3)-ABX4(AN1)
Y4=ABY4(AN3)-ABY4(AN1)
Z4=ABZ4(AN3)-ABZ4(AN1)
->350
300:X2=XC3-ABX4(AN1)
    Y2=YC3-ABY4(AN1)
    Z2=ZC3-ABZ4(AN1)
X3=ABX4(AN2)-ABX4(AN1)
Y3=ABY4(AN2)-ABY4(AN1)
Z3=ABZ4(AN2)-ABZ4(AN1)
X4=XC4-ABX4(AN1)
Y4=YC4-ABY4(AN1)
Z4=ZC4-ABZ4(AN1)
350:RR=SQRT(X2↑2+Y2↑2)
CO=X2/RR;SI=Y2/RR
DHM1(1,1)= CO;DHM1(1,2)= SI;DHM1(1,3)=  0
DHM1(2,1)=-SI;DHM1(2,2)= CO;DHM1(2,3)=  0
DHM1(3,1)=  0;DHM1(3,2)=  0;DHM1(3,3)=  1
RRR=SQRT(X2↑2+Y2↑2+Z2↑2)
CO=RR/RRR
SI=Z2/RRR
DHM2(1,1)= CO;DHM2(1,2)=  0;DHM2(1,3)= SI
DHM2(2,1)=  0;DHM2(2,2)=  1;DHM2(2,3)=  0
DHM2(3,1)=-SI;DHM2(3,2)=  0;DHM2(3,3)= CO
MATRIXPROD(DHM3,DHM2,DHM1)
AX'=DHM3(1,1)*X3+DHM3(1,2)*Y3+DHM3(1,3)*Z3
AY'=DHM3(2,1)*X3+DHM3(2,2)*Y3+DHM3(2,3)*Z3
AZ'=DHM3(3,1)*X3+DHM3(3,2)*Y3+DHM3(3,3)*Z3
RR=SQRT(AY'↑2+AZ'↑2)
CO=-AZ'/RR
SI=-AY'/RR
DHM1(1,1)=  1;DHM1(1,2)=  0;DHM1(1,3)=  0
DHM1(2,1)=  0;DHM1(2,2)= CO;DHM1(2,3)=-SI
DHM1(3,1)=  0;DHM1(3,2)= SI;DHM1(3,3)= CO
MATRIXPROD(DHM,DHM1,DHM3)
```

```
S1=DHM(2,1)*X4+DHM(2,2)*Y4+DHM(2,3)*Z4
S2=DHM(3,1)*X4+DHM(3,2)*Y4+DHM(3,3)*Z4
RR=SQRT(S1↑2+S2↑2)
CO=S2/RR
SI=-S1/RR
RRR=QUADRANT(HAND*SI,CO)
->205 %IF J'=2
%CAPTION ~DIHEDRAL␣ANGLE␣PHI2=
->210
205:%CAPTION ~DIHEDRAL␣ANGLE␣PSI2=
210:PRINT(((180*RRR)/π),3,2)
%REPEAT

%BEGIN
%REAL CO,SI,ZA
%INTEGER I,J
%ARRAY X,Y,Z(1:2*(NA+NB))
%INTEGER %ARRAY ATOMTYPE(1:2*(NA+NB))

%CYCLE I=-1,1,0
CO=COS(I*2*π/NN)
SI=SIN(I*2*π/NN)
ZA=I*HR/NN
%CYCLE J=1,1,NA
X((I+1)*(NA+NB)+J)=ABX4(J)*CO - ABY4(J)*SI
Y((I+1)*(NA+NB)+J)=ABY4(J)*CO + ABX4(J)*SI
Z((I+1)*(NA+NB)+J)=ABZ4(J) + ZA
ATOMTYPE((I+1)*(NA+NB)+J)=AT(J)
%REPEAT
%CYCLE J=(NA+1),1,(NA+NB)
X((I+1)*(NA+NB)+J)=BX4(J-NA)*CO - BY4(J-NA)*SI
Y((I+1)*(NA+NB)+J)=BY4(J-NA)*CO + BX4(J-NA)*SI
Z((I+1)*(NA+NB)+J)=BZ4(J-NA) + ZA
ATOMTYPE((I+1)*(NA+NB)+J)=AT'(J-NA)
%REPEAT
%REPEAT
%CYCLE I=1,1,NA
%CYCLE J=(NA+1),1,(NA+NB)
->140 %IF I=AN4 %OR J=(BN4+NA)
->140 %IF I=AN5 %AND J=(BN5+NA)
->160 %IF (X(I)-X(J))↑2 + (Y(I)-Y(J))↑2 + (Z(I)-Z(J))↑2 <%C
VDW(ATOMTYPE(I),ATOMTYPE(J))
140:%REPEAT;%REPEAT
%CAPTION ~CONFORMATION␣AT␣B-A␣LINKAGE␣IS␣ALLOWED
->180
160:%CAPTION ~CONFORMATION␣AT␣B-A␣LINKAGE␣IS␣NOT␣ALLOWED
180:%CYCLE I=1,1,NA
%CYCLE J=((2*NA)+NB+1),1,(2*(NA+NB))
->340 %IF I=(AN1) %OR J=(BN1+(2*NA)+NB)
->340 %IF I=(AN2) %AND J=(BN2+(2*NA)+NB)
->360 %IF (X(I)-X(J))↑2 + (Y(I)-Y(J))↑2 + (Z(I)-Z(J))↑2 < %C
VDW(ATOMTYPE(I),ATOMTYPE(J))
340:%REPEAT;%REPEAT
%CAPTION ~CONFORMATION␣AT␣A-B␣LINKAGE␣IS␣ALLOWED
->380
360:%CAPTION ~CONFORMATION␣AT␣A-B␣LINKAGE␣IS␣NOT␣ALLOWED
380:%END
```

```
%REPEAT;->200
100:%CAPTION ~LENGTH␣OF␣DISACCHARIDE␣IS␣LESS␣THAN␣(1/NN)*HELIX␣REPEAT
->200
175:%CAPTION ~DISCRIMINANT␣OF␣REDUCING␣CUBIC␣IS␣NEGATIVE
200:%REPEAT
    %REPEAT

%ROUTINE MATRIXPROD(%REALARRAYNAME C,A,B)
%CYCLE I=1,1,3;%CYCLE J=1,1,3
C(I,J)=A(I,1)*B(1,J)+A(I,2)*B(2,J)+A(I,3)*B(3,J)
%REPEAT;%REPEAT
%END

%REALFN QUADRANT(%REAL SINANG,COSANG)
ANG=ARCSIN(!(SINANG)!)
->5 %IF SINANG>=0 %AND COSANG>=0
->6 %IF SINANG>=0 %AND COSANG <0
->7 %IF SINANG <0 %AND COSANG>=0
->8 %IF SINANG <0 %AND COSANG <0
5:VANG=ANG  ;->9
6:VANG=(π-ANG)  ;->9
7:VANG=(2*π)-ANG  ;->9
8:VANG=π+ANG  ;->9
9:%RESULT=VANG
%END

%END %OF %PROGRAM
```

```
  1    4·0674    0·0000    1·0719    1
  2    3·4705   -1·2385    0·3986    1
  3    1·9683   -1·2701    0·6641    1
  4    1·4332    0·0000    0·0000    1
  5    1·9719    1·1436    0·8643    1
  6    1·8295    0·4806    2·2523    1
  7    5·3321    0·0990    0·5140    2
  8    3·6770   -1·1532   -1·0172    2
  9    1·8600   -0·9266    2·0472    2
 10    0·0000    0·0000    0·0000    2
 11    3·3777    1·1700    0·6156    2
 12    4·0903   -0·0790    2·1487    3
 13    3·9353   -2·1327    0·7876    3
 14    1·4699   -2·1838    0·3774    3
 15    1·8090    0·0873   -1·0087    3
 16    1·6783    2·1742    0·9994    3
 17    0·7562    0·7716    2·6005    3
 18    2·5674    0·8078    2·9773    3
-1

7   1
```

```
 1   -1.3901    0.0000    0.0000    1
 2   -1.9047   -1.2563    0.6910    1
 3   -3.4339   -1.2331    0.8137    1
 4   -3.8374    0.0000    1.5465    1
 5   -3.3103    1.1947    0.7702    1
 6   -3.6778    2.5019    1.3954    1
 7    0.0000    0.0000    0.0000    2
 8   -1.5357   -2.4323   -0.0548    2
 9   -3.8561   -2.3870    1.5544    2
10   -3.3654    0.1217    2.8690    2
11   -1.8696    1.1559    0.6731    2
12   -1.7545   -0.0202   -1.0153    3
13   -1.4736   -1.3166    1.6795    3
14   -3.8827   -1.2415   -0.1690    3
15   -4.9160    0.0509    1.5625    3
16   -3.7223    1.1643   -0.2277    3
-1

9  3

116

-30   10   70
-50   10  -20

26.0

3.0

116

  4     11
10      7
  4     1

3

3.0   2.7   2.2
2.7   2.7   2.2
2.2   2.2   1.9

***Z
```

DESCRIPTION OF THE DATA INPUT TO MODELBUILD I.

1	AX1(1)	AY1(1)	AZ1(1)	AT(1)	The number of atom, x, y z coordinates and atom type of each of the NA atoms of residue A w.r.t the axes AX, AY, AZ
2	AX1(2)	AY1(2)	AZ1(2)	AT(2)	
.	.	.	.	.	
.	.	.	.	.	
NA	AX1(NA)	AY1(NA)	AZ1(NA)	AT(NA)	

-1 — Atom list terminated by -1.

AN1 — AN1 = the number in the above list of the oxygen atom at the A-B linkage. (e.g. O(1) of 3,6-anhydrogalactose.)

AN2 — AN2 = the number of the carbon atom in the above list to which the oxygen atom AN1 is attached. (e.g. C(1) of 3,6-anhydrogalactose.)

1	BX1(1)	BY1(1)	BZ1(1)	AT'(1)	The number of atom, x, y z coordinates and atom type of each of the NB atoms of residue B w.r.t the axes BX, BY, BZ.
2	BX1(2)	BY1(2)	BZ1(2)	AT'(2)	
.	.	.	.	.	
.	.	.	.	.	
NB	BX1(NB)	BY1(NB)	BZ1(NB)	AT'(NB)	

-1 — Atom list terminated by -1.

BN1 — BN1 = the number in the above list of the oxygen atom at the A-B linkage. (e.g. O(3) of galactose.)

BN2 — BN2 = the number of the carbon atom in the above list to which the oxygen atom BN1 is attached. (e.g. C(3) of galactose.)

ALPHA' — The bond angle (in degrees) at the bridge oxygen atom of the B-A linkage. (e.g. at O(1/4) of galactose/3,6-anhydro-galactose.)

PHI1 DPHI PHI2 PSI1 DPSI PSI2	Ranges of ϕ and ψ to be covered; minimum values, intervals and maximum values in degrees. (All 6 parameters must be integers and the ranges must be an integral number of intervals.)
HR	Pitch of the helix.
NN	The screw symmetry; positive if a right handed helix and negative if a left handed helix.
OMEGA	Bond angle required at the bridge oxygen atom of the A-B linkage. (e.g. O(1/3) of the 3,6-anhydro-galactose/galactose.)
AN3	The number of the atom in the list of coord-inates of A which def-ines the zero of ϕ_2. (e.g. C(4) of 3,6-anhydro-galactose.)
BN3	The number of the atom in the list of coord-inates of B which def-ines the zero of ψ_2.(e.g. O(5) of galactose.)
AN4	The number of the atom in the list of coord-inates of A which def-ines the bridge oxygen atom at the B-A linkage. (e.g. O(4) of 3,6-anhydrogalactose.)
BN4	The number of the atom in the list of coord-inates of B which def-ines the bridge oxygen atom at the B-A linkage. (e.g. O(1) of galactose.)

AN5	The number in the list of coordinates of A of the carbon atom to which the oxygen atom AN4 is attached. (e.g. C(4) of 3,6-anhydro-galactose.)
BN5	The number in the list of coordinates of B of the carbon atom to which the oxygen atom BN4 is attached. (e.g. C(1) of galactose.)
NAT	The number of atom types.
VDW'(1,1) . . . . VDW'(1,NAT) . . . . VDW'(NAT,1) . . . VDW'(NAT,NAT)	Table of van der Waals contact distances VDW'(i,j) between atom types i and j. (This gives a table of values NAT×NAT.)
***Z	Terminates the data.

Note. The atom type AT(i) or AT'(i) is an integer between 1 and NAT.

APPENDIX II

TABLES OF OBSERVED AND CALCULATED STRUCTURE FACTORS

(ANGAL DIFFRACTOMETER DATA)

Angal structure factor tables. (based on diffractometer data.)

Columns are h,k,l, 10.468.$|F|_{obs.}$ and 10.468.$|F|_{calc.}$

H	K	L	Y(OBS)	Y(CALC)
0	0	2	232.00	224.17
0	0	4	108.00	106.45
0	0	6	180.00	186.06
0	0	8	39.00	20.23
0	1	1	168.00	145.61
0	1	2	34.00	36.02
0	1	3	58.00	65.46
0	1	4	41.00	14.31
0	1	5	58.00	47.73
0	1	6	13.00	7.72
0	1	7	74.00	73.65
0	1	8	36.00	28.42
0	2	0	847.00	951.05
0	2	1	462.00	417.78
0	2	2	79.00	72.31
0	2	3	301.00	263.75
0	2	4	193.00	180.65
0	2	5	277.00	277.56
0	2	6	49.00	43.58
0	2	7	33.00	26.98
0	2	8	31.00	23.49
0	3	1	184.00	165.61
0	3	2	347.00	303.69
0	3	3	288.00	256.01
0	3	4	227.00	216.48
0	3	5	202.00	196.90
0	3	6	88.00	88.56
0	3	7	98.00	103.86
0	3	8	33.00	37.68
0	4	0	668.00	612.18
0	4	1	72.00	33.44
0	4	2	92.00	89.71
0	4	3	147.00	143.26
0	4	4	108.00	99.18
0	4	5	47.00	41.43
0	4	6	51.00	45.53
0	4	7	16.00	8.30
0	5	1	356.00	340.05
0	5	2	75.00	58.65
0	5	3	168.00	161.50
0	5	4	174.00	166.72
0	5	5	44.00	48.27
0	5	6	141.00	157.54
0	5	7	36.00	42.92
0	6	0	220.00	189.21
0	6	1	103.00	116.90
0	6	2	272.00	264.30
0	6	3	92.00	103.03
0	6	4	128.00	125.93
0	6	5	33.00	28.46
0	6	6	144.00	144.10
0	6	7	30.00	28.31
0	7	1	281.00	274.91
0	7	2	174.00	180.13
0	7	3	36.00	43.04
0	7	4	92.00	83.38
0	7	5	51.00	56.44
0	7	6	44.00	46.50
0	7	7	116.00	113.63
0	8	0	100.00	88.30
0	8	1	250.00	234.20
0	8	2	51.00	52.87
0	8	3	276.00	272.84
0	8	4	33.00	33.55
0	8	5	13.00	2.94
0	8	6	151.00	151.21
0	9	1	53.00	49.60
0	9	2	22.00	19.21
0	9	3	152.00	157.99
0	9	4	152.00	153.12
0	9	5	24.00	23.91
0	9	6	21.00	16.38
0	10	0	28.00	22.15
0	10	1	41.00	46.22
0	10	2	78.00	77.55
0	10	3	185.00	192.09
0	10	4	114.00	122.07
0	10	5	31.00	16.66
0	11	0	18.00	0.00
0	11	1	72.00	71.16
0	11	2	58.00	62.94
0	11	3	15.00	4.38
0	11	4	55.00	59.37
0	11	5	26.00	20.74
0	12	0	85.00	90.14
0	12	1	33.00	35.28
0	12	2	40.00	31.48
0	12	3	15.00	7.25
0	12	4	29.00	7.42
0	13	1	22.00	17.15
0	13	2	62.00	53.06
0	13	3	70.00	62.39
0	14	0	92.00	91.73
0	14	1	36.00	39.89
1	0	1	532.00	545.24
1	0	2	461.00	493.27
1	0	3	55.00	41.36

1	0	4	55.00	55.51	1	8	3	231.00	243.64
1	0	5	99.00	102.64	1	8	4	27.00	21.25
1	0	6	222.00	240.89	1	8	5	80.00	81.31
1	0	7	88.00	94.37	1	8	6	66.00	57.74
1	0	8	19.00	13.97	1	9	0	222.00	233.54
1	1	1	283.00	260.18	1	9	1	99.00	99.95
1	1	2	598.00	554.08	1	9	2	82.00	79.33
1	1	3	294.00	276.82	1	9	3	226.00	232.66
1	1	4	224.00	214.53	1	9	4	55.00	48.66
1	1	5	56.00	62.43	1	9	5	49.00	39.46
1	1	6	166.00	166.69	1	9	6	26.00	24.31
1	1	7	99.00	100.10	1	10	1	121.00	124.69
1	1	8	14.00	4.23	1	10	2	149.00	154.51
1	2	0	364.00	371.50	1	10	3	100.00	96.26
1	2	1	483.00	435.18	1	10	4	66.00	67.98
1	2	2	13.00	15.87	1	10	5	18.00	11.00
1	2	3	322.00	305.18	1	11	1	53.00	45.77
1	2	4	182.00	178.00	1	11	2	135.00	135.97
1	2	5	170.00	170.67	1	11	3	62.00	61.53
1	2	6	90.00	95.55	1	11	4	60.00	62.55
1	2	7	160.00	160.84	1	11	5	20.00	18.07
1	2	8	17.00	4.31	1	12	0	62.00	67.35
1	3	0	114.00	156.62	1	12	1	17.00	13.18
1	3	1	270.00	250.99	1	12	2	31.00	32.20
1	3	2	133.00	131.68	1	12	3	24.00	16.88
1	3	3	476.00	439.63	1	12	4	8.00	7.57
1	3	4	93.00	87.96	1	13	0	16.00	3.18
1	3	5	76.00	76.76	1	13	1	35.00	31.45
1	3	6	51.00	51.64	1	13	2	30.00	20.57
1	3	7	73.00	72.24	1	13	3	52.00	46.21
1	3	8	32.00	32.03	1	14	1	16.00	12.70
1	4	0	155.00	169.06	2	0	0	81.00	73.92
1	4	1	102.00	82.45	2	0	1	597.00	546.26
1	4	2	122.00	108.74	2	0	2	33.00	27.64
1	4	3	374.00	349.12	2	0	3	302.00	284.43
1	4	4	55.00	49.97	2	0	4	58.00	60.33
1	4	5	67.00	68.85	2	0	5	53.00	52.48
1	4	6	84.00	80.83	2	0	6	100.00	96.42
1	4	7	59.00	54.75	2	0	7	52.00	45.97
1	5	0	812.00	791.94	2	0	8	17.00	17.01
1	5	1	179.00	164.93	2	1	0	1206.00	1219.13
1	5	2	113.00	101.33	2	1	1	173.00	133.72
1	5	3	225.00	211.57	2	1	2	176.00	162.73
1	5	4	68.00	69.95	2	1	3	280.00	266.28
1	5	5	22.00	24.71	2	1	4	75.00	62.69
1	5	6	60.00	72.34	2	1	5	138.00	142.57
1	5	7	94.00	97.79	2	1	6	68.00	61.94
1	6	0	474.00	456.62	2	1	7	87.00	92.30
1	6	1	164.00	156.55	2	1	8	45.00	45.69
1	6	2	80.00	81.52	2	2	0	340.00	331.39
1	6	3	25.00	29.60	2	2	1	376.00	372.78
1	6	4	58.00	55.80	2	2	2	393.00	365.28
1	6	5	15.00	16.45	2	2	3	88.00	87.17
1	6	6	27.00	28.92	2	2	4	112.00	109.62
1	6	7	115.00	118.96	2	2	5	153.00	148.95
1	7	0	175.00	178.66	2	2	6	60.00	54.41
1	7	1	74.00	74.64	2	2	7	51.00	52.84
1	7	2	69.00	74.37	2	2	8	54.00	51.30
1	7	3	131.00	132.02	2	3	0	292.00	250.57
1	7	4	168.00	171.65	2	3	1	237.00	214.33
1	7	5	152.00	156.50	2	3	2	93.00	86.73
1	7	6	65.00	65.05	2	3	3	70.00	65.21
1	7	7	82.00	81.50	2	3	4	84.00	78.17
1	8	0	71.00	80.55	2	3	5	122.00	118.27
1	8	1	71.00	67.50	2	3	6	62.00	63.16
1	8	2	144.00	141.19	2	3	7	53.00	55.39

2	4	0	257.00	253.91	2	13	3	73.00	72.16
2	4	1	223.00	217.92	2	14	0	19.00	2.50
2	4	2	226.00	220.43	3	0	1	229.00	220.35
2	4	3	180.00	158.10	3	0	2	127.00	125.92
2	4	4	60.00	59.04	3	0	3	314.00	302.83
2	4	5	141.00	141.98	3	0	4	26.00	26.55
2	4	6	105.00	114.52	3	0	5	187.00	183.43
2	4	7	42.00	47.70	3	0	6	112.00	109.95
2	5	0	41.00	60.48	3	0	7	9.00	1.80
2	5	1	119.00	118.98	3	1	0	285.00	281.48
2	5	2	91.00	83.48	3	1	1	300.00	292.03
2	5	3	203.00	187.96	3	1	2	217.00	203.65
2	5	4	27.00	28.13	3	1	3	112.00	121.33
2	5	5	40.00	32.27	3	1	4	96.00	96.54
2	5	6	25.00	31.62	3	1	5	313.00	311.29
2	5	7	39.00	39.84	3	1	6	151.00	162.93
2	6	0	119.00	113.11	3	1	7	82.00	78.28
2	6	1	281.00	266.41	3	2	0	126.00	126.76
2	6	2	241.00	233.77	3	2	1	252.00	236.67
2	6	3	131.00	128.58	3	2	2	408.00	388.54
2	6	4	52.00	55.04	3	2	3	64.00	63.00
2	6	5	159.00	165.04	3	2	4	99.00	101.36
2	6	6	105.00	113.52	3	2	5	183.00	186.67
2	6	7	90.00	87.12	3	2	6	93.00	89.94
2	7	0	23.00	18.59	3	2	7	27.00	26.17
2	7	1	78.00	71.98	3	3	0	196.00	193.82
2	7	2	103.00	98.15	3	3	1	130.00	125.51
2	7	3	198.00	196.42	3	3	2	161.00	148.50
2	7	4	46.00	52.48	3	3	3	162.00	139.52
2	7	5	108.00	110.71	3	3	4	147.00	146.48
2	7	6	97.00	94.28	3	3	5	87.00	86.40
2	7	7	66.00	63.85	3	3	6	68.00	68.06
2	8	0	13.00	5.46	3	3	7	34.00	30.18
2	8	1	131.00	138.39	3	4	0	446.00	448.62
2	8	2	87.00	91.63	3	4	1	133.00	136.10
2	8	3	186.00	189.64	3	4	2	111.00	110.35
2	8	4	100.00	94.57	3	4	3	212.00	204.01
2	8	5	34.00	37.28	3	4	4	61.00	57.94
2	8	6	21.00	20.09	3	4	5	53.00	60.39
2	9	0	39.00	47.93	3	4	6	72.00	74.46
2	9	1	91.00	91.83	3	4	7	43.00	45.04
2	9	2	69.00	65.90	3	5	0	78.00	78.57
2	9	3	247.00	254.72	3	5	1	144.00	142.75
2	9	4	58.00	62.48	3	5	2	122.00	113.87
2	9	5	17.00	13.61	3	5	3	77.00	72.61
2	9	6	19.00	16.08	3	5	4	149.00	162.23
2	10	0	63.00	58.35	3	5	5	110.00	120.38
2	10	1	50.00	50.47	3	5	6	54.00	54.14
2	10	2	123.00	123.37	3	5	7	73.00	72.19
2	10	3	28.00	20.35	3	6	0	307.00	300.51
2	10	4	66.00	71.43	3	6	1	108.00	117.03
2	10	5	25.00	26.55	3	6	2	72.00	63.43
2	11	0	7.00	10.92	3	6	3	30.00	22.28
2	11	1	44.00	51.81	3	6	4	111.00	119.39
2	11	2	24.00	24.04	3	6	5	123.00	128.98
2	11	3	50.00	47.03	3	6	6	79.00	83.15
2	11	4	54.00	50.84	3	6	7	78.00	80.14
2	11	5	21.00	23.79	3	7	0	95.00	95.57
2	12	0	19.00	7.62	3	7	1	228.00	229.81
2	12	1	80.00	86.57	3	7	2	136.00	136.40
2	12	2	54.00	49.86	3	7	3	94.00	97.56
2	12	3	62.00	54.46	3	7	4	124.00	124.27
2	12	4	38.00	34.53	3	7	5	55.00	53.63
2	13	0	33.00	31.58	3	7	6	24.00	21.35
2	13	1	21.00	19.32	3	8	0	219.00	227.93
2	13	2	25.00	19.98	3	8	1	54.00	53.02

3	8	2	113.00	115.51	4	4	4	151.00	150.51
3	8	3	28.00	33.78	4	4	5	126.00	139.42
3	8	4	71.00	64.65	4	4	6	72.00	79.62
3	8	5	133.00	135.83	4	4	7	97.00	99.61
3	8	6	20.00	10.61	4	5	0	150.00	157.37
3	9	0	20.00	1.26	4	5	1	86.00	92.08
3	9	1	130.00	132.41	4	5	2	144.00	145.27
3	9	2	193.00	204.58	4	5	3	145.00	141.20
3	9	3	116.00	115.75	4	5	4	109.00	107.55
3	9	4	99.00	102.72	4	5	5	143.00	145.19
3	9	5	35.00	30.04	4	5	6	59.00	62.94
3	9	6	16.00	6.47	4	5	7	55.00	55.26
3	10	0	140.00	140.50	4	6	0	57.00	54.30
3	10	1	68.00	81.06	4	6	1	63.00	69.31
3	10	2	32.00	32.71	4	6	2	147.00	149.41
3	10	3	63.00	63.71	4	6	3	12.00	2.77
3	10	4	59.00	57.96	4	6	4	56.00	52.45
3	10	5	37.00	41.02	4	6	5	46.00	44.77
3	11	0	21.00	25.06	4	6	6	29.00	24.69
3	11	1	100.00	99.47	4	7	0	52.00	40.13
3	11	2	15.00	6.62	4	7	1	49.00	53.26
3	11	3	18.00	4.05	4	7	2	264.00	273.24
3	11	4	40.00	36.61	4	7	3	135.00	144.99
3	12	0	8.00	6.45	4	7	4	69.00	72.01
3	12	1	16.00	22.89	4	7	5	116.00	122.10
3	12	2	59.00	57.32	4	7	6	43.00	41.63
3	12	3	40.00	32.74	4	8	0	132.00	131.91
3	13	0	27.00	21.73	4	8	1	59.00	56.10
3	13	1	45.00	43.12	4	8	2	113.00	113.28
3	13	2	64.00	62.03	4	8	3	153.00	159.66
4	0	0	67.00	67.54	4	8	4	42.00	40.85
4	0	1	334.00	330.30	4	8	5	44.00	46.63
4	0	2	108.00	102.21	4	8	6	49.00	49.02
4	0	3	56.00	57.65	4	9	0	40.00	29.79
4	0	4	157.00	146.80	4	9	1	102.00	107.96
4	0	5	166.00	175.01	4	9	2	114.00	118.18
4	0	6	49.00	42.05	4	9	3	17.00	18.28
4	0	7	126.00	127.78	4	9	4	72.00	73.16
4	1	0	490.00	486.00	4	9	5	44.00	39.56
4	1	1	342.00	340.97	4	10	1	51.00	48.18
4	1	2	212.00	218.17	4	10	2	45.00	46.25
4	1	3	47.00	45.46	4	10	3	69.00	68.28
4	1	4	247.00	249.95	4	10	4	58.00	59.57
4	1	5	174.00	177.75	4	10	5	46.00	43.04
4	1	6	135.00	130.58	4	11	0	68.00	73.18
4	1	7	10.00	10.39	4	11	1	43.00	48.21
4	2	0	34.00	15.71	4	11	2	85.00	86.25
4	2	1	138.00	133.02	4	11	3	55.00	50.21
4	2	2	341.00	315.87	4	11	4	17.00	15.49
4	2	3	88.00	90.08	4	12	0	14.00	11.41
4	2	4	160.00	151.86	4	12	1	47.00	52.19
4	2	5	62.00	59.01	4	12	2	53.00	50.85
4	2	6	80.00	81.22	4	12	3	45.00	45.77
4	2	7	41.00	45.50	4	13	0	69.00	66.57
4	3	0	198.00	203.90	4	13	1	58.00	54.80
4	3	1	46.00	38.11	5	0	1	203.00	206.26
4	3	2	363.00	346.41	5	0	2	406.00	406.11
4	3	3	80.00	76.81	5	0	3	60.00	59.99
4	3	4	91.00	88.08	5	0	4	44.00	37.18
4	3	5	123.00	122.44	5	0	5	233.00	226.87
4	3	6	71.00	74.57	5	0	6	30.00	31.56
4	3	7	57.00	57.76	5	0	7	10.00	4.40
4	4	0	40.00	18.75	5	1	0	116.00	129.33
4	4	1	15.00	15.22	5	1	1	77.00	77.20
4	4	2	171.00	165.98	5	1	2	114.00	105.25
4	4	3	57.00	52.91	5	1	3	83.00	77.96

5	1	4	152.00	148.93	5	11	1	81.00	89.88
5	1	5	178.00	184.67	5	11	2	31.00	27.23
5	1	6	32.00	29.57	5	11	3	42.00	42.31
5	1	7	44.00	41.74	5	12	1	42.00	39.05
5	2	0	203.00	212.17	5	12	2	74.00	69.91
5	2	1	204.00	204.01	6	0	0	209.00	208.40
5	2	2	208.00	202.75	6	0	1	70.00	75.37
5	2	3	206.00	204.12	6	0	2	31.00	33.18
5	2	4	29.00	21.38	6	0	3	75.00	69.21
5	2	5	100.00	100.72	6	0	4	166.00	171.05
5	2	6	52.00	52.92	6	0	5	31.00	25.17
5	2	7	39.00	33.52	6	0	6	46.00	21.52
5	3	0	55.00	46.88	6	1	0	92.00	93.01
5	3	1	174.00	171.87	6	1	1	146.00	148.14
5	3	2	31.00	28.36	6	1	2	87.00	80.54
5	3	3	108.00	113.27	6	1	3	123.00	123.82
5	3	4	93.00	96.03	6	1	4	116.00	120.05
5	3	5	102.00	108.99	6	1	5	36.00	36.16
5	3	6	40.00	36.67	6	1	6	41.00	40.13
5	3	7	46.00	41.31	6	2	0	67.00	64.72
5	4	0	166.00	167.72	6	2	1	130.00	131.27
5	4	1	148.00	141.21	6	2	2	99.00	102.98
5	4	2	92.00	76.72	6	2	3	25.00	23.27
5	4	3	118.00	122.04	6	2	4	102.00	101.17
5	4	4	62.00	55.10	6	2	5	83.00	85.09
5	4	5	108.00	111.13	6	2	6	75.00	77.04
5	4	6	64.00	65.20	6	3	0	120.00	122.36
5	5	0	26.00	21.85	6	3	1	196.00	193.88
5	5	1	102.00	106.18	6	3	2	61.00	54.26
5	5	2	58.00	50.79	6	3	3	49.00	43.03
5	5	3	96.00	91.03	6	3	4	79.00	85.32
5	5	4	71.00	78.13	6	3	5	100.00	101.71
5	5	5	49.00	45.46	6	3	6	46.00	48.11
5	5	6	15.00	5.96	6	4	0	22.00	25.07
5	6	0	103.00	108.73	6	4	1	70.00	66.89
5	6	1	112.00	120.97	6	4	2	36.00	31.75
5	6	2	47.00	54.97	6	4	3	41.00	39.30
5	6	3	18.00	22.44	6	4	4	21.00	14.44
5	6	4	52.00	52.31	6	4	5	57.00	59.15
5	6	5	37.00	36.11	6	4	6	18.00	12.97
5	6	6	8.00	3.91	6	5	1	154.00	159.77
5	7	0	21.00	32.21	6	5	2	94.00	102.15
5	7	1	80.00	80.27	6	5	3	60.00	65.44
5	7	2	153.00	160.24	6	5	4	60.00	67.14
5	7	3	96.00	105.11	6	5	5	80.00	82.97
5	7	4	81.00	81.10	6	5	6	21.00	8.95
5	7	5	55.00	52.64	6	6	0	86.00	84.49
5	7	6	37.00	34.72	6	6	1	99.00	101.44
5	8	0	54.00	56.42	6	6	2	117.00	119.24
5	8	1	154.00	147.61	6	6	3	47.00	42.29
5	8	2	116.00	122.13	6	6	4	39.00	39.06
5	8	3	63.00	60.41	6	6	5	112.00	112.29
5	8	4	99.00	103.65	6	6	6	23.00	12.85
5	8	5	21.00	17.22	6	7	0	155.00	162.90
5	9	0	68.00	65.19	6	7	1	64.00	64.09
5	9	1	87.00	92.84	6	7	2	90.00	92.30
5	9	2	131.00	133.10	6	7	3	118.00	125.02
5	9	3	42.00	39.74	6	7	4	36.00	41.59
5	9	4	38.00	33.87	6	7	5	21.00	13.61
5	9	5	19.00	11.81	6	8	0	50.00	48.82
5	10	0	40.00	34.61	6	8	1	43.00	46.52
5	10	1	120.00	121.44	6	8	2	91.00	85.25
5	10	2	34.00	38.60	6	8	3	38.00	38.86
5	10	3	80.00	75.41	6	8	4	30.00	31.40
5	10	4	33.00	35.95	6	8	5	41.00	43.18
5	11	0	14.00	11.33	6	9	0	20.00	17.55

6	9	1	95.00	98.16	7	8	2	107.00	105.75
6	9	2	37.00	41.83	7	8	3	75.00	74.55
6	9	3	81.00	78.34	7	8	4	44.00	42.37
6	9	4	46.00	42.45	7	9	0	50.00	52.29
6	10	0	59.00	61.52	7	9	1	88.00	92.99
6	10	1	14.00	7.61	7	9	2	32.00	35.52
6	10	2	22.00	16.74	7	9	3	53.00	49.82
6	10	3	76.00	83.17	7	10	0	25.00	29.93
6	11	0	50.00	47.11	7	10	1	46.00	47.60
6	11	1	66.00	66.04	7	10	2	39.00	37.60
6	11	2	60.00	60.13	7	11	1	19.00	13.74
6	12	0	52.00	44.51	8	0	0	253.00	242.50
7	0	1	176.00	181.11	8	0	1	15.00	10.37
7	0	2	16.00	15.37	8	0	2	34.00	32.98
7	0	3	25.00	28.06	8	0	3	55.00	61.14
7	0	4	123.00	125.89	8	0	4	55.00	55.78
7	0	5	62.00	63.55	8	0	5	30.00	33.26
7	0	6	40.00	39.76	8	1	0	20.00	19.38
7	1	0	75.00	78.50	8	1	1	27.00	27.28
7	1	1	94.00	92.59	8	1	2	91.00	93.90
7	1	2	118.00	115.44	8	1	3	96.00	98.32
7	1	3	46.00	43.01	8	1	4	72.00	71.10
7	1	4	85.00	89.51	8	1	5	29.00	27.59
7	1	5	56.00	52.06	8	2	0	58.00	60.25
7	1	6	44.00	40.39	8	2	1	83.00	88.35
7	2	0	256.00	268.89	8	2	2	82.00	85.21
7	2	1	52.00	53.74	8	2	3	59.00	67.52
7	2	2	17.00	7.24	8	2	4	33.00	33.77
7	2	3	7.00	1.20	8	2	5	17.00	11.34
7	2	4	60.00	63.92	8	3	0	29.00	40.11
7	2	5	45.00	44.79	8	3	1	66.00	65.46
7	2	6	58.00	56.79	8	3	2	125.00	129.86
7	3	0	23.00	36.03	8	3	3	38.00	33.76
7	3	1	30.00	33.41	8	3	4	80.00	76.10
7	3	2	82.00	85.44	8	3	5	51.00	53.09
7	3	3	102.00	113.29	8	4	0	157.00	167.32
7	3	4	49.00	51.02	8	4	1	100.00	100.71
7	3	5	21.00	19.27	8	4	2	78.00	81.50
7	3	6	38.00	37.51	8	4	3	44.00	38.57
7	4	0	37.00	35.76	8	4	4	40.00	39.26
7	4	1	117.00	120.04	8	4	5	19.00	12.44
7	4	2	44.00	47.98	8	5	0	63.00	70.88
7	4	3	41.00	35.56	8	5	1	24.00	21.48
7	4	4	70.00	69.97	8	5	2	52.00	56.36
7	4	5	32.00	28.30	8	5	3	42.00	46.18
7	4	6	16.00	10.20	8	5	4	63.00	66.62
7	5	0	157.00	161.36	8	5	5	38.00	33.85
7	5	1	16.00	6.00	8	6	0	67.00	66.16
7	5	2	83.00	85.62	8	6	1	77.00	78.26
7	5	3	76.00	75.39	8	6	2	53.00	55.39
7	5	4	18.00	17.58	8	6	3	82.00	87.19
7	5	5	18.00	17.22	8	6	4	35.00	33.47
7	6	0	59.00	59.31	8	7	0	40.00	35.69
7	6	1	96.00	96.97	8	7	1	22.00	18.38
7	6	2	107.00	114.91	8	7	2	29.00	25.78
7	6	3	81.00	83.40	8	7	3	44.00	43.33
7	6	4	67.00	71.74	8	7	4	40.00	40.91
7	6	5	26.00	26.22	8	8	0	18.00	7.06
7	7	0	28.00	21.54	8	8	1	27.00	22.61
7	7	1	82.00	89.51	8	8	2	33.00	32.73
7	7	2	39.00	34.38	8	8	3	53.00	50.33
7	7	3	75.00	76.84	8	9	0	43.00	42.11
7	7	4	43.00	40.21	8	9	1	24.00	15.04
7	7	5	17.00	18.48	8	9	2	55.00	51.88
7	8	0	55.00	53.52	8	10	0	30.00	27.28
7	8	1	67.00	64.76	9	0	1	120.00	121.63

9	0	2	50.00	51.32	0	1	10	89.00	94.28
9	0	3	40.00	33.90	0	1	11	26.00	16.38
9	0	4	37.00	32.68	0	2	9	75.00	73.62
9	1	0	88.00	91.81	0	2	10	36.00	32.16
9	1	1	92.00	89.85	0	2	11	23.00	19.64
9	1	2	41.00	41.97	0	3	9	25.00	3.98
9	1	3	31.00	36.02	0	3	10	94.00	99.61
9	1	4	28.00	25.33	0	3	11	17.00	21.09
9	2	0	81.00	85.68	0	4	8	19.00	12.47
9	2	1	61.00	62.11	0	4	9	55.00	57.32
9	2	2	36.00	39.04	0	4	10	9.00	1.15
9	2	3	20.00	19.62	0	5	8	37.00	38.42
9	2	4	34.00	31.31	0	5	9	58.00	59.12
9	3	0	47.00	49.51	0	5	10	36.00	31.55
9	3	1	18.00	17.15	0	6	8	19.00	15.37
9	3	2	19.00	14.76	0	6	9	29.00	26.13
9	3	3	58.00	58.00	0	6	10	19.00	6.23
9	3	4	59.00	54.21	0	7	8	33.00	27.20
9	4	0	34.00	32.99	0	7	9	17.00	5.27
9	4	1	41.00	37.47	0	7	10	36.00	36.33
9	4	2	53.00	50.62	0	8	7	88.00	86.71
9	4	3	59.00	62.06	0	8	8	28.00	8.51
9	4	4	17.00	14.71	0	8	9	75.00	74.56
9	5	0	68.00	67.05	0	8	10	21.00	12.15
9	5	1	62.00	61.68	0	9	7	9.00	0.89
9	5	2	22.00	18.22	0	9	8	38.00	35.80
9	5	3	24.00	17.26	0	9	9	51.00	43.50
9	6	0	28.00	30.05	0	10	6	64.00	61.70
9	6	1	44.00	47.81	0	10	7	9.00	5.57
9	6	2	35.00	30.75	0	10	8	19.00	8.11
9	6	3	47.00	42.92	0	10	9	33.00	29.66
9	7	0	41.00	38.05	0	11	6	19.00	14.64
9	7	1	59.00	57.93	0	11	7	58.00	50.17
9	7	2	15.00	12.15	0	11	8	39.00	36.64
9	8	0	20.00	26.54	0	11	9	19.00	15.74
9	8	1	53.00	49.87	0	12	5	16.00	13.35
10	0	0	42.00	6.63	0	12	6	71.00	65.90
10	0	1	48.00	51.38	0	12	7	60.00	68.63
10	0	2	60.00	58.37	0	12	8	27.00	22.56
10	0	3	60.00	60.82	0	13	4	80.00	80.82
10	1	0	72.00	71.05	0	13	5	17.00	1.46
10	1	1	22.00	14.76	0	13	6	18.00	15.29
10	1	2	24.00	20.04	0	13	7	17.00	2.35
10	1	3	35.00	41.54	0	13	8	21.00	7.20
10	2	1	63.00	63.62	0	14	2	27.00	31.05
10	2	2	44.00	38.63	0	14	3	18.00	9.62
10	2	3	35.00	33.03	0	14	4	20.00	15.56
10	3	1	46.00	47.02	0	14	5	19.00	16.55
10	3	2	30.00	24.52	0	14	6	22.00	17.91
10	3	3	20.00	14.38	0	14	7	20.00	11.51
10	4	0	19.00	10.26	0	15	0	29.00	0.00
10	4	1	33.00	35.01	0	15	1	53.00	56.58
10	4	2	26.00	21.49	0	15	2	52.00	51.56
10	5	0	40.00	25.67	0	15	3	75.00	79.61
10	5	1	43.00	39.95	0	15	4	35.00	38.19
10	5	2	15.00	4.88	0	15	5	29.00	25.33
10	6	0	32.00	38.72	0	15	6	17.00	7.91
10	6	1	17.00	11.78	0	15	7	22.00	11.41
11	0	1	27.00	15.02	0	16	0	26.00	8.52
11	1	1	19.00	6.10	0	16	1	15.00	5.22
11	2	0	32.00	33.59	0	16	2	63.00	60.01
11	2	1	50.00	46.80	0	16	3	19.00	11.22
0	0	9	24.00	0.00	0	16	4	25.00	23.21
0	0	10	72.00	63.29	0	16	5	25.00	12.61
0	0	11	18.00	0.00	0	16	6	20.00	9.36
0	1	9	25.00	21.63	0	17	0	21.00	0.00

0	17	1	40.00	42.12	1	14	7	20.00	15.08
0	17	2	47.00	43.78	1	15	0	24.00	23.84
0	17	3	24.00	8.77	1	15	1	49.00	47.66
0	17	4	21.00	12.86	1	15	2	56.00	52.88
0	17	5	15.00	3.06	1	15	3	54.00	54.06
0	18	0	61.00	60.91	1	15	4	67.00	64.22
0	18	1	23.00	22.93	1	15	5	23.00	17.96
0	18	2	15.00	4.26	1	15	6	21.00	12.41
0	18	3	15.00	5.90	1	15	7	20.00	7.36
0	18	4	14.00	4.32	1	16	0	19.00	8.46
0	19	0	20.00	0.00	1	16	1	43.00	33.41
0	19	1	42.00	37.58	1	16	2	20.00	15.96
0	19	2	24.00	1.16	1	16	3	30.00	29.79
1	0	9	25.00	16.79	1	16	4	57.00	56.24
1	0	10	43.00	39.35	1	16	5	15.00	1.89
1	0	11	15.00	0.77	1	16	6	20.00	20.18
1	1	9	36.00	33.17	1	17	0	63.00	65.73
1	1	10	53.00	51.52	1	17	1	31.00	29.79
1	1	11	19.00	7.95	1	17	2	23.00	27.05
1	2	9	23.00	18.24	1	17	3	35.00	33.16
1	2	10	39.00	37.96	1	17	4	33.00	33.17
1	2	11	17.00	6.93	1	17	5	20.00	16.13
1	3	9	49.00	46.37	1	18	0	23.00	14.05
1	3	10	28.00	29.34	1	18	1	24.00	20.00
1	3	11	15.00	15.97	1	18	2	22.00	14.02
1	4	8	17.00	9.64	1	18	3	21.00	14.60
1	4	9	33.00	30.65	1	18	4	26.00	19.77
1	4	10	58.00	60.83	1	19	0	23.00	12.96
1	5	8	25.00	20.29	1	19	1	23.00	7.56
1	5	9	21.00	14.26	1	19	2	14.00	6.33
1	5	10	23.00	17.61	2	0	9	32.00	25.45
1	6	8	33.00	36.02	2	0	10	25.00	22.42
1	6	9	18.00	10.80	2	0	11	25.00	14.69
1	6	10	44.00	43.07	2	1	9	57.00	57.43
1	7	8	27.00	23.89	2	1	10	40.00	38.11
1	7	9	44.00	44.07	2	1	11	31.00	27.90
1	7	10	36.00	38.31	2	2	9	21.00	13.80
1	8	7	44.00	41.02	2	2	10	77.00	78.49
1	8	8	28.00	25.99	2	2	11	23.00	19.95
1	8	9	28.00	29.18	2	3	8	22.00	21.92
1	8	10	32.00	29.43	2	3	9	55.00	52.38
1	9	7	17.00	17.26	2	3	10	20.00	17.66
1	9	8	34.00	35.16	2	4	8	71.00	66.36
1	9	9	39.00	37.16	2	4	9	31.00	31.30
1	10	6	25.00	24.40					
1	10	7	18.00	16.16					
1	10	8	49.00	51.51					
1	10	9	15.00	12.82	2	4	10	21.00	16.35
1	11	6	20.00	15.51	2	5	8	29.00	28.03
1	11	7	9.00	6.59	2	5	9	19.00	13.67
1	11	8	39.00	38.92	2	5	10	18.00	8.31
1	11	9	22.00	15.79	2	6	8	49.00	49.85
1	12	5	24.00	23.01	2	6	9	25.00	23.57
1	12	6	50.00	49.67	2	6	10	35.00	31.14
1	12	7	32.00	29.19	2	7	8	51.00	50.59
1	12	8	17.00	14.99	2	7	9	56.00	56.27
1	13	4	43.00	35.96	2	7	10	25.00	22.34
1	13	5	15.00	21.52	2	8	7	87.00	86.76
1	13	6	54.00	56.97	2	8	8	15.00	20.90
1	13	7	24.00	17.90	2	8	9	35.00	32.31
1	13	8	13.00	2.46	2	8	10	13.00	11.02
1	14	2	36.00	31.32	2	9	7	29.00	21.37
1	14	3	61.00	64.88	2	9	8	53.00	50.62
1	14	4	79.00	77.84	2	9	9	51.00	51.37
1	14	5	18.00	10.56	2	10	6	13.00	7.10
1	14	6	44.00	45.06	2	10	7	83.00	82.05

2	10	8	53.00	50.26	3	6	6	63.00	57.39
2	10	9	17.00	10.96	3	6	9	46.00	47.40
2	11	6	67.00	63.03	3	6	10	22.00	14.22
2	11	7	49.00	52.50	3	7	7	29.00	29.33
2	11	8	36.00	30.74	3	7	8	45.00	42.14
2	11	9	25.00	23.46	3	7	9	21.00	5.87
2	12	5	23.00	16.28	3	7	10	33.00	30.37
2	12	6	15.00	5.51	3	8	7	16.00	5.65
2	12	7	39.00	35.10	3	8	8	29.00	27.90
2	12	8	23.00	19.62	3	8	9	68.00	66.58
2	13	4	20.00	15.93	3	9	7	39.00	37.20
2	13	5	19.00	19.27	3	9	8	42.00	39.61
2	13	6	12.00	9.42	3	9	9	21.00	19.07
2	13	7	25.00	29.27	3	10	6	19.00	17.13
2	13	8	26.00	26.01	3	10	7	17.00	11.68
2	14	1	44.00	48.19	3	10	8	60.00	61.89
2	14	2	46.00	39.25	3	10	9	21.00	14.03
2	14	3	69.00	71.32	3	11	5	30.00	22.39
2	14	4	41.00	41.63	3	11	6	51.00	46.80
2	14	5	20.00	13.72	3	11	7	33.00	27.15
2	14	6	25.00	12.94	3	11	8	38.00	36.31
2	14	7	29.00	25.05	3	12	4	27.00	22.98
2	15	0	33.00	36.65	3	12	5	31.00	31.18
2	15	1	68.00	68.97	3	12	6	31.00	30.28
2	15	2	52.00	49.57	3	12	7	25.00	19.70
2	15	3	14.00	5.96	3	12	8	18.00	18.18
2	15	4	39.00	38.79	3	13	3	57.00	59.20
2	15	5	29.00	28.50	3	13	4	53.00	55.75
2	15	6	29.00	26.68	3	13	5	17.00	5.16
2	16	0	15.00	7.24	3	13	6	32.00	35.47
2	16	1	49.00	49.68	3	13	7	21.00	22.88
2	16	2	53.00	56.76	3	14	0	13.00	2.80
2	16	3	38.00	36.29	3	14	1	75.00	73.27
2	16	4	21.00	18.57	3	14	2	29.00	30.75
2	16	5	15.00	12.32	3	14	3	55.00	52.05
2	16	6	25.00	23.54	3	14	4	45.00	34.42
2	17	0	40.00	43.50	3	14	5	23.00	18.26
2	17	1	22.00	12.08	3	14	6	18.00	3.48
2	17	2	26.00	20.25	3	14	7	21.00	13.64
2	17	3	19.00	19.72	3	15	0	27.00	22.69
2	17	4	29.00	21.88	3	15	1	17.00	7.05
2	17	5	24.00	18.81	3	15	2	19.00	16.28
2	18	0	26.00	24.12	3	15	3	41.00	39.74
2	18	1	55.00	54.42	3	15	4	41.00	39.15
2	18	2	24.00	14.80	3	15	5	20.00	9.37
2	18	3	17.00	17.81	3	15	6	12.00	6.91
2	19	0	32.00	23.75	3	16	0	74.00	79.73
2	19	1	29.00	26.82	3	16	1	26.00	15.64
3	0	8	23.00	17.60	3	16	2	44.00	42.00
3	0	9	24.00	22.27	3	16	3	24.00	26.93
3	0	10	13.00	7.74	3	16	4	32.00	30.23
3	1	8	19.00	13.28	3	16	5	30.00	21.75
3	1	9	35.00	31.88	3	17	0	32.00	23.98
3	1	10	36.00	37.33	3	17	1	17.00	13.30
3	2	8	70.00	69.55	3	17	2	25.00	20.05
3	2	9	26.00	23.61	3	17	3	22.00	18.50
3	2	10	23.00	13.46	3	17	4	36.00	31.97
3	3	8	46.00	40.37	3	18	0	40.00	37.66
3	3	9	12.00	6.67	3	18	1	34.00	29.65
3	3	10	71.00	75.55	3	18	2	25.00	19.67
3	4	8	25.00	24.55	3	18	3	20.00	13.97
3	4	9	41.00	46.59	3	19	0	18.00	5.37
3	4	10	20.00	16.31	4	0	8	51.00	49.93
3	5	8	12.00	6.82	4	0	9	19.00	15.09
3	5	9	32.00	25.67	4	0	10	16.00	8.42
3	5	10	30.00	29.45	4	1	8	31.00	30.00

4	1	9	33.00	26.50	4	16	5	26.00	22.48
4	1	10	15.00	4.66	4	17	0	22.00	16.32
4	2	8	53.00	57.30	4	17	1	55.00	53.15
4	2	9	34.00	27.03	4	17	2	29.00	26.79
4	2	10	45.00	49.95	4	17	3	9.00	5.83
4	3	8	51.00	50.76	4	17	4	25.00	21.10
4	3	9	25.00	21.41	4	18	0	31.00	30.55
4	3	10	18.00	10.49	4	18	1	15.00	11.10
4	4	8	72.00	72.22	4	18	2	9.00	6.72
4	4	9	16.00	7.12	5	0	8	57.00	54.99
4	4	10	22.00	14.91	5	0	9	33.00	27.36
4	5	8	78.00	80.72	5	0	10	40.00	43.60
4	5	9	33.00	29.75	5	1	8	43.00	44.12
4	5	10	22.00	18.44	5	1	9	30.00	27.04
4	6	7	15.00	10.86	5	1	10	16.00	8.85
4	6	8	15.00	16.60	5	2	8	42.00	37.80
4	6	9	49.00	50.70	5	2	9	19.00	10.93
4	6	10	17.00	8.45	5	2	10	43.00	43.12
4	7	7	67.00	64.12	5	3	8	63.00	66.93
4	7	8	25.00	22.75	5	3	9	13.00	13.89
4	7	9	19.00	13.87	5	3	10	30.00	33.97
4	8	7	28.00	21.64	5	4	7	23.00	24.69
4	8	8	98.00	93.47	5	4	8	19.00	12.61
4	8	9	30.00	26.88	5	4	9	9.00	7.67
4	9	6	14.00	7.35	5	4	10	32.00	29.54
4	9	7	59.00	55.23	5	5	7	68.00	67.36
4	9	8	28.00	29.55	5	5	8	27.00	24.97
4	9	9	15.00	15.20	5	5	9	36.00	40.32
4	10	6	33.00	33.48	5	5	10	19.00	8.73
4	10	7	19.00	12.70	5	6	7	42.00	46.36
4	10	8	50.00	44.97	5	6	8	52.00	50.99
4	10	9	40.00	38.69	5	6	9	22.00	19.43
4	11	5	28.00	21.01	5	7	7	14.00	13.09
4	11	6	28.00	24.14	5	7	8	27.00	24.77
4	11	7	63.00	62.29	5	7	9	59.00	58.03
4	11	8	21.00	15.72	5	8	6	23.00	24.38
4	12	4	34.00	25.18	5	8	7	43.00	44.28
4	12	5	24.00	22.00	5	8	8	43.00	37.90
4	12	6	16.00	11.24	5	8	9	27.00	24.74
4	12	7	28.00	29.98	5	9	6	16.00	12.90
4	12	8	35.00	28.96	5	9	7	29.00	24.11
4	13	2	19.00	8.27	5	9	8	38.00	38.96
4	13	3	24.00	24.04	5	9	9	25.00	21.66
4	13	4	18.00	10.54	5	10	5	25.00	19.21
4	13	5	21.00	11.74	5	10	6	36.00	32.98
4	13	6	12.00	7.23	5	10	7	58.00	55.81
4	13	7	13.00	2.57	5	10	8	40.00	41.62
4	14	1	50.00	52.71	5	11	4	36.00	38.11
4	14	2	26.00	22.05	5	11	5	36.00	32.95
4	14	3	24.00	23.22	5	11	6	26.00	22.54
4	14	4	33.00	30.93	5	11	7	24.00	19.43
4	14	5	33.00	31.57	5	11	8	16.00	12.20
4	14	6	17.00	15.44	5	12	3	45.00	45.50
4	14	7	21.00	4.24	5	12	4	15.00	10.69
4	15	0	22.00	20.99	5	12	5	18.00	10.07
4	15	1	33.00	36.05	5	12	6	47.00	47.15
4	15	2	24.00	22.19	5	12	7	15.00	13.63
4	15	3	32.00	32.70	5	13	0	21.00	3.33
4	15	4	21.00	21.60	5	13	1	37.00	37.24
4	15	5	24.00	13.42	5	13	2	57.00	57.14
4	15	6	28.00	24.36	5	13	3	31.00	27.23
4	16	0	16.00	1.63	5	13	4	34.00	27.27
4	16	1	51.00	48.81	5	13	5	19.00	14.58
4	16	2	17.00	9.73	5	13	6	19.00	19.95
4	16	3	23.00	20.08	5	13	7	21.00	19.76
4	16	4	26.00	20.95	5	14	0	34.00	36.15

5	14	1	11.00	9.74	6	10	6	13.00	11.77
5	14	2	29.00	28.99	6	10	7	56.00	54.34
5	14	3	29.00	33.20	6	10	8	25.00	19.67
5	14	4	24.00	16.68	6	11	3	32.00	30.50
5	14	5	17.00	9.38	6	11	4	23.00	15.96
5	14	6	20.00	19.92	6	11	5	43.00	40.09
5	15	0	39.00	42.25	6	11	6	26.00	20.82
5	15	1	23.00	23.73	6	11	7	29.00	27.64
5	15	2	28.00	31.53	6	11	8	13.00	3.66
5	15	3	35.00	29.75	6	12	1	33.00	31.50
5	15	4	41.00	36.89	6	12	2	13.00	8.22
5	15	5	29.00	28.02	6	12	3	33.00	35.79
5	15	6	25.00	15.82	6	12	4	22.00	21.95
5	16	0	58.00	51.43	6	12	5	19.00	12.75
5	16	1	24.00	22.72	6	12	6	16.00	9.23
5	16	2	12.00	3.72	6	12	7	16.00	17.28
5	16	3	21.00	14.93	6	13	0	15.00	3.46
5	16	4	16.00	20.29	6	13	1	23.00	25.31
5	16	5	29.00	22.75	6	13	2	53.00	46.93
5	17	0	51.00	46.93	6	13	3	33.00	33.83
5	17	1	32.00	28.82	6	13	4	44.00	41.37
5	17	2	24.00	16.39	6	13	5	26.00	22.78
5	17	3	24.00	19.19	6	13	6	18.00	10.44
5	18	0	37.00	34.38	6	13	7	19.00	17.87
5	18	1	24.00	10.54	6	14	0	33.00	30.21
5	18	2	31.00	25.04	6	14	1	29.00	31.59
6	0	7	80.00	79.65	6	14	2	37.00	36.61
6	0	8	23.00	17.86	6	14	3	44.00	38.49
6	0	9	15.00	1.45	6	14	4	25.00	20.45
6	0	10	13.00	1.67	6	14	5	27.00	26.28
6	1	7	45.00	37.31	6	14	6	27.00	22.45
6	1	8	36.00	34.34	6	15	0	17.00	15.83
6	1	9	12.00	3.67	6	15	1	44.00	42.52
6	1	10	15.00	10.84	6	15	2	27.00	19.75
6	2	7	24.00	23.15	6	15	3	15.00	5.61
6	2	8	13.00	7.11	6	15	4	26.00	20.66
6	2	9	21.00	17.10	6	15	5	24.00	18.36
6	2	10	39.00	39.63	6	16	0	28.00	19.74
6	3	7	18.00	11.27	6	16	1	48.00	44.46
6	3	8	29.00	27.81	6	16	2	17.00	13.44
6	3	9	30.00	25.59	6	16	3	20.00	14.53
6	3	10	17.00	10.23	6	16	4	19.00	11.53
6	4	7	60.00	58.96	6	17	0	9.00	8.68
6	4	8	33.00	31.95	6	17	1	28.00	28.26
6	4	9	9.00	12.75	6	17	2	10.00	15.69
6	4	10	21.00	12.58	6	17	3	24.00	18.50
6	5	7	28.00	21.45	7	0	7	55.00	46.04
6	5	8	22.00	18.90	7	0	8	16.00	10.05
6	5	9	19.00	8.49	7	0	9	31.00	28.29
6	6	7	14.00	8.18	7	1	7	14.00	17.02
6	6	8	20.00	7.08	7	1	8	23.00	16.64
6	6	9	24.00	10.06	7	1	9	24.00	18.41
6	7	6	35.00	32.18	7	2	7	55.00	48.99
6	7	7	38.00	35.62	7	2	8	34.00	31.26
6	7	8	46.00	41.82	7	2	9	28.00	22.20
6	7	9	21.00	17.95	7	3	7	14.00	6.33
6	8	6	21.00	14.63	7	3	8	24.00	15.11
6	8	7	47.00	47.31	7	3	9	42.00	41.95
6	8	8	23.00	17.51	7	4	7	37.00	35.07
6	8	9	27.00	24.16	7	4	8	19.00	15.51
6	9	5	72.00	72.36	7	4	9	16.00	3.21
6	9	6	25.00	20.70	7	5	6	21.00	4.80
6	9	7	19.00	12.07	7	5	7	50.00	53.19
6	9	8	37.00	31.27	7	5	8	20.00	22.09
6	10	4	41.00	35.23	7	5	9	17.00	16.47
6	10	5	19.00	13.04	7	6	6	14.00	1.64

7	5	7	58.00	56.13	8	1	5	25.00	21.62
7	6	6	20.00	14.51	8	1	9	27.00	17.33
7	6	9	40.00	33.83	8	2	6	62.00	56.63
7	7	6	27.00	20.20	8	2	7	16.00	7.64
7	7	7	28.00	22.97	8	2	8	20.00	18.11
7	7	8	27.00	19.17	8	2	9	23.00	10.90
7	7	9	30.00	23.76	8	3	6	27.00	19.19
7	8	5	27.00	23.33	8	3	7	22.00	16.07
7	8	6	30.00	31.04	8	3	8	33.00	27.72
7	8	7	32.00	23.72	8	3	9	27.00	24.61
7	8	8	29.00	26.78	8	4	6	17.00	6.15
7	9	4	36.00	26.68	8	4	7	31.00	31.51
7	9	5	21.00	17.48	8	4	8	37.00	34.95
7	9	6	20.00	16.01	8	4	9	21.00	10.77
7	9	7	33.00	32.53	8	5	6	40.00	37.62
7	9	8	30.00	30.93	8	5	7	16.00	7.42
7	10	3	17.00	11.74	8	5	8	13.00	6.72
7	10	4	35.00	32.15	8	5	9	30.00	26.98
7	10	5	38.00	33.91	8	6	5	31.00	32.91
7	10	6	42.00	39.55	8	6	6	39.00	37.91
7	10	7	32.00	30.64	8	6	7	22.00	18.24
7	10	8	13.00	16.61	8	6	8	46.00	45.40
7	11	2	56.00	50.21	8	7	5	40.00	36.03
7	11	3	43.00	40.71	8	7	6	31.00	24.18
7	11	4	27.00	23.39	8	7	7	43.00	42.53
7	11	5	27.00	23.10	8	7	8	21.00	15.38
7	11	6	30.00	28.84	8	8	4	37.00	34.87
7	11	7	22.00	15.00	8	8	5	17.00	17.03
7	12	0	28.00	26.86	8	8	6	25.00	18.68
7	12	1	11.00	4.13	8	8	7	43.00	40.83
7	12	2	45.00	38.32	8	8	8	23.00	19.28
7	12	3	19.00	8.94	8	9	3	58.00	56.38
7	12	4	41.00	36.95	8	9	4	103.00	98.17
7	12	5	33.00	21.49	8	9	5	27.00	21.23
7	12	6	27.00	25.27	8	9	6	24.00	19.94
7	12	7	17.00	3.50	8	9	7	34.00	29.87
7	13	0	35.00	27.51	8	10	1	41.00	37.70
7	13	1	23.00	16.65	8	10	2	25.00	20.30
7	13	2	26.00	20.46	8	10	3	40.00	41.48
7	13	3	26.00	23.92	8	10	4	28.00	25.72
7	13	4	21.30	18.22	8	10	5	20.00	13.36
7	13	5	20.00	8.18	8	10	6	33.00	29.64
7	13	6	27.00	25.46	8	10	7	14.00	1.59
7	14	0	50.00	41.93	8	11	0	19.00	17.86
7	14	1	16.00	15.63	8	11	1	21.00	22.75
7	14	2	49.00	42.61	8	11	2	27.00	22.47
7	14	3	27.00	21.33	8	11	3	46.00	41.17
7	14	4	35.00	35.98	8	11	4	41.00	32.60
7	14	5	14.00	15.27	8	11	5	19.00	13.80
7	15	0	44.00	37.55	8	11	6	25.00	12.71
7	15	1	18.00	13.55	8	11	7	18.00	2.80
7	15	2	28.00	24.16	8	12	0	9.00	11.48
7	15	3	23.00	15.09	8	12	1	34.00	32.97
7	15	4	41.00	34.76	8	12	2	46.00	41.03
7	16	0	53.00	45.08	8	12	3	27.00	25.15
7	16	1	25.00	23.61	8	12	4	21.00	12.97
7	16	2	20.00	12.61	8	12	5	17.00	9.34
7	16	3	27.00	18.60	8	12	6	36.00	34.34
7	17	0	36.00	25.40	8	13	0	26.00	9.51
7	17	1	25.00	23.57	8	13	1	28.00	28.37
8	0	6	66.00	61.30	8	13	2	16.00	15.83
8	0	7	19.00	17.15	8	13	3	22.00	17.11
8	0	8	24.00	16.10	8	13	4	19.00	21.84
8	0	9	9.00	0.65	8	13	5	21.00	19.78
8	1	6	55.00	49.33	8	14	0	21.00	10.36
8	1	7	85.00	82.05	8	14	1	28.00	20.77

8	14	2	31.00	23.23	9	10	4	32.00	29.64
8	14	3	24.00	21.03	9	10	5	24.00	17.21
8	14	4	24.00	14.90	9	10	6	33.00	29.65
8	14	5	15.00	7.04	9	11	0	19.00	16.86
8	15	0	34.00	24.20	9	11	1	26.00	20.56
8	15	1	28.00	24.99	9	11	2	32.00	27.01
8	15	2	17.00	5.03	9	11	3	43.00	38.22
8	15	3	22.00	11.51	9	11	4	42.00	36.66
8	16	0	17.00	9.19	9	11	5	27.00	21.83
8	16	1	18.00	17.02	9	11	6	23.00	16.89
8	16	2	28.00	19.99	9	12	0	27.00	21.35
9	0	5	45.00	39.04	9	12	1	18.00	15.17
9	0	6	51.00	42.16	9	12	2	21.00	20.86
9	0	7	30.00	18.75	9	12	3	20.00	10.29
9	0	8	23.00	15.19	9	12	4	25.00	19.01
9	0	9	34.00	31.69	9	12	5	21.00	16.36
9	1	5	52.00	50.84	9	13	0	36.00	27.14
9	1	6	59.00	57.75	9	13	1	9.00	8.22
9	1	7	43.00	37.41	9	13	2	9.00	7.36
9	1	8	25.00	16.05	9	13	3	18.00	10.31
9	2	5	65.00	61.49	9	13	4	15.00	3.52
9	2	6	28.00	21.62	9	13	5	22.00	8.07
9	2	7	9.00	12.27	9	14	0	22.00	12.18
9	2	8	14.00	5.07	9	14	1	27.00	22.27
9	3	5	38.00	34.09	9	14	2	25.00	24.82
9	3	6	29.00	25.69	9	14	3	23.00	14.23
9	3	7	22.00	16.91	9	14	4	34.00	31.99
9	3	8	21.00	14.74	9	15	0	23.00	16.69
9	4	5	44.00	45.48	9	15	1	23.00	18.10
9	4	6	29.00	31.14	9	15	2	23.00	15.40
9	4	7	31.00	30.26	10	0	4	32.00	25.70
9	4	8	14.00	9.81	10	0	5	41.00	40.15
9	5	4	37.00	36.38	10	0	6	20.00	12.13
9	5	5	41.00	36.40	10	0	7	49.00	45.52
9	5	6	21.00	16.50	10	0	8	30.00	27.57
9	5	7	35.00	35.06	10	1	4	22.00	14.63
9	5	8	24.00	19.58	10	1	5	12.00	9.90
9	6	4	16.00	10.56	10	1	6	21.00	13.96
9	6	5	16.00	12.15	10	1	7	9.00	10.01
9	6	6	25.00	22.89	10	1	8	14.00	2.99
9	6	7	47.00	48.06	10	2	4	17.00	13.22
9	6	8	15.00	11.70	10	2	5	39.00	35.52
9	7	3	44.00	47.34	10	2	6	18.00	13.63
9	7	4	18.00	21.41	10	2	7	29.00	27.62
9	7	5	11.00	7.76	10	2	8	25.00	22.01
9	7	6	24.00	20.30	10	3	4	15.00	15.25
9	7	7	29.00	28.89	10	3	5	35.00	32.59
9	7	8	22.00	17.54	10	3	6	31.00	25.47
9	8	2	57.00	56.48	10	3	7	25.00	24.00
9	8	3	68.00	63.96	10	3	8	26.00	17.56
9	8	4	26.00	18.30	10	4	3	23.00	16.67
9	8	5	13.00	6.19	10	4	4	28.00	25.94
9	8	6	35.00	34.69	10	4	5	18.00	10.21
9	8	7	29.00	25.42	10	4	6	32.00	30.41
9	9	0	24.00	17.07	10	4	7	21.00	17.00
9	9	1	21.00	22.99	10	4	8	15.00	4.53
9	9	2	16.00	10.70	10	5	3	24.00	13.17
9	9	3	51.00	43.54	10	5	4	38.00	32.57
9	9	4	26.00	16.61	10	5	5	23.00	13.55
9	9	5	26.00	20.92	10	5	6	18.00	7.81
9	9	6	37.00	31.68	10	5	7	21.00	16.15
9	9	7	21.00	7.56	10	6	2	45.00	41.42
9	10	0	18.00	13.77	10	6	3	22.00	14.85
9	10	1	18.00	10.30	10	6	4	21.00	17.57
9	10	2	31.00	25.53	10	6	5	31.00	29.01
9	10	3	21.00	16.43	10	6	6	31.00	25.01

10	6	7	31.00	25.97	11	3	1	14.00	0.77
10	7	1	38.00	27.08	11	3	2	15.00	12.84
10	7	2	28.00	25.98	11	3	3	21.00	20.06
10	7	3	20.00	20.00	11	3	4	25.00	24.79
10	7	4	36.00	32.57	11	3	5	44.00	36.72
10	7	5	34.00	29.54	11	3	6	21.00	14.32
10	7	6	20.00	14.83	11	3	7	23.00	16.96
10	7	7	26.00	18.35	11	4	0	67.00	63.15
10	8	0	20.00	13.11	11	4	1	24.00	23.85
10	8	1	16.00	9.00	11	4	2	27.00	27.24
10	8	2	19.00	17.05	11	4	3	35.00	36.25
10	8	3	14.00	10.90	11	4	4	37.00	32.71
10	8	4	57.00	52.99	11	4	5	32.00	24.67
10	8	5	31.00	23.91	11	4	6	25.00	21.30
10	8	6	31.00	24.90	11	4	7	23.00	19.15
10	8	7	22.00	11.66	11	5	0	8.00	2.92
10	9	0	15.00	0.76	11	5	1	26.00	23.09
10	9	1	24.00	20.73	11	5	2	45.00	38.56
10	9	2	17.00	7.11	11	5	3	28.00	25.97
10	9	3	41.00	39.35	11	5	4	22.00	17.51
10	9	4	24.00	20.47	11	5	5	41.00	38.65
10	9	5	19.00	13.42	11	5	6	18.00	18.25
10	9	6	14.00	10.16	11	5	7	19.00	13.55
10	10	0	21.00	12.98	11	6	0	37.00	27.40
10	10	1	34.00	32.45	11	6	1	23.00	14.62
10	10	2	28.00	22.44	11	6	2	34.00	32.16
10	10	3	21.00	15.78	11	6	3	23.00	15.49
10	10	4	42.00	35.95	11	6	4	20.00	11.01
10	10	5	36.00	34.07	11	6	5	16.00	11.03
10	10	6	23.00	19.63	11	6	6	18.00	8.81
10	11	0	9.00	5.51	11	7	0	15.00	6.62
10	11	1	34.00	24.92	11	7	1	34.00	29.19
10	11	2	27.00	24.38	11	7	2	65.00	61.39
10	11	3	18.00	16.78	11	7	3	32.00	24.00
10	11	4	16.00	5.79	11	7	4	30.00	23.95
10	11	5	18.00	1.78	11	7	5	23.00	13.97
10	12	0	9.00	7.91	11	7	6	17.00	3.96
10	12	1	34.00	28.62	11	8	0	25.00	6.91
10	12	2	25.00	13.59	11	8	1	12.00	11.25
10	12	3	10.00	4.55	11	8	2	24.00	12.03
10	12	4	24.00	15.67	11	8	3	18.00	13.24
10	13	0	9.00	6.79	11	8	4	14.00	5.68
10	13	1	21.00	12.57	11	8	5	15.00	10.64
10	13	2	21.00	11.43	11	8	6	21.00	11.08
10	13	3	43.00	38.09	11	9	0	9.00	3.51
10	14	0	24.00	19.93	11	9	1	17.00	1.66
10	14	1	20.00	9.97	11	9	2	46.00	40.44
10	14	2	17.00	10.68	11	9	3	16.00	8.80
11	0	2	19.00	18.03	11	9	4	15.00	3.56
11	0	3	28.00	24.97	11	9	5	10.00	6.92
11	0	4	17.00	5.57	11	10	1	30.00	21.06
11	0	5	19.00	13.94	11	10	2	19.00	15.55
11	0	6	35.00	35.29	11	10	3	24.00	21.83
11	0	7	23.00	12.35	11	10	4	13.00	8.92
11	1	2	23.00	24.99	11	10	5	17.00	10.64
11	1	3	13.00	13.75	11	11	0	21.00	15.88
11	1	4	31.00	29.42	11	11	1	16.00	10.38
11	1	5	31.00	21.87	11	11	2	18.00	14.64
11	1	6	24.00	17.43	11	11	3	20.00	6.67
11	1	7	19.00	10.10	11	11	4	31.00	22.88
11	2	2	22.00	22.31	11	12	0	21.00	6.19
11	2	3	34.00	27.38	11	12	1	17.00	6.40
11	2	4	34.00	27.41	11	12	2	20.00	15.15
11	2	5	40.00	38.87	11	12	3	19.00	12.31
11	2	6	35.00	28.85	11	13	0	24.00	17.16
11	2	7	24.00	17.37	11	13	1	20.00	5.18

12	0	0	23.00	4.94	12	10	1	20.00	15.30
12	0	1	41.00	37.73	12	10	2	20.00	12.09
12	0	2	31.00	29.89	12	10	3	11.00	7.73
12	0	3	21.00	10.63	12	11	0	21.00	6.61
12	0	4	11.00	0.63	12	11	1	23.00	13.27
12	0	5	19.00	11.43	12	11	2	24.00	20.95
12	0	6	28.00	24.09	13	0	0	16.00	0.00
12	1	0	38.00	29.57	13	0	1	23.00	15.55
12	1	1	19.00	16.95	13	0	2	21.00	25.66
12	1	2	18.00	12.49	13	0	3	24.00	22.08
12	1	3	19.00	9.80	13	0	4	13.00	3.00
12	1	4	23.00	22.36	13	0	5	38.00	31.11
12	1	5	38.00	36.11	13	1	0	17.00	11.69
12	1	6	24.00	18.47	13	1	1	35.00	28.15
12	2	0	34.00	31.82	13	1	2	14.00	6.22
12	2	1	47.00	43.27	13	1	3	32.00	23.63
12	2	2	56.00	50.45	13	1	4	23.00	18.86
12	2	3	17.00	3.93	13	1	5	27.00	21.01
12	2	4	33.00	29.17	13	2	0	21.00	24.57
12	2	5	23.00	17.58	13	2	1	14.00	3.45
12	2	6	10.00	12.34	13	2	2	20.00	19.62
12	3	0	17.00	14.74	13	2	3	25.00	19.07
12	3	1	28.00	24.51	13	2	4	12.00	10.90
12	3	2	24.00	22.80	13	2	5	24.00	20.37
12	3	3	16.00	11.35	13	3	0	27.00	14.72
12	3	4	28.00	24.92	13	3	1	32.00	28.25
12	3	5	29.00	25.11	13	3	2	27.00	17.60
12	3	6	23.00	14.80	13	3	3	37.00	31.22
12	4	0	61.00	58.17	13	3	4	25.00	19.71
12	4	1	34.00	33.08	13	3	5	20.00	15.58
12	4	2	31.00	26.91	13	4	0	23.00	10.52
12	4	3	26.00	19.78	13	4	1	31.00	26.19
12	4	4	23.00	21.19	13	4	2	18.00	8.45
12	4	5	21.00	19.07	13	4	3	23.00	10.34
12	4	6	18.00	10.36	13	4	4	17.00	11.51
12	5	0	18.00	4.40	13	4	5	24.00	22.34
12	5	1	27.00	25.47	13	5	0	15.00	2.02
12	5	2	34.00	30.74	13	5	1	18.00	5.31
12	5	3	35.00	33.59	13	5	2	33.00	29.92
12	5	4	30.00	22.81	13	5	3	25.00	19.71
12	5	5	17.00	7.74	13	5	4	19.00	6.11
12	5	6	20.00	17.67	13	5	5	16.00	10.53
12	6	0	45.00	39.15	13	6	0	17.00	0.51
12	6	1	12.00	9.60	13	6	1	32.00	28.90
12	6	2	21.00	16.53	13	6	2	25.00	17.13
12	6	3	25.00	20.46	13	6	3	21.00	12.63
12	6	4	37.00	35.07	13	6	4	14.00	5.03
12	6	5	10.00	4.18	13	7	0	19.00	9.66
12	7	0	21.00	9.07	13	7	1	25.00	19.90
12	7	1	23.00	15.96	13	7	2	21.00	12.50
12	7	2	41.00	37.05	13	7	3	20.00	17.42
12	7	3	24.00	17.43	13	7	4	20.00	10.30
12	7	4	14.00	4.82	13	8	0	9.00	4.47
12	7	5	10.00	8.60	13	8	1	25.00	18.72
12	8	0	14.00	15.01	13	8	2	30.00	22.14
12	8	1	14.00	9.89	13	8	3	16.00	5.23
12	8	2	33.00	23.95	13	9	0	28.00	25.61
12	8	3	24.00	20.22	13	9	1	17.00	11.62
12	8	4	25.00	21.33	13	9	2	24.00	6.77
12	8	5	21.00	15.70	13	10	0	21.00	16.72
12	9	0	10.00	2.76	14	0	1	12.00	13.98
12	9	1	29.00	22.31	14	0	2	23.00	16.80
12	9	2	13.00	0.43	14	0	3	27.00	25.81
12	9	3	18.00	13.15	14	0	4	17.00	13.24
12	9	4	16.00	8.42	14	1	0	21.00	15.13
12	10	0	17.00	3.05	14	1	1	9.00	9.62

14	1	2	21.00	13.36
14	1	3	19.00	3.55
14	1	4	13.00	12.50
14	2	1	20.00	17.87
14	2	2	23.00	12.69
14	2	3	20.00	15.72
14	2	4	15.00	5.20
14	3	0	44.00	37.15
14	3	1	14.00	3.94
14	3	2	19.00	11.84
14	3	3	18.00	7.44
14	3	4	18.00	7.00
14	4	0	16.00	0.99
14	4	1	19.00	8.48
14	4	2	14.00	1.09
14	4	3	16.00	6.91
14	5	0	26.00	21.96
14	5	1	15.00	10.56
14	5	2	20.00	3.73
14	5	3	29.00	18.70
14	6	0	19.00	6.00
14	6	1	10.00	4.52
14	6	2	19.00	8.02
14	7	0	18.00	0.84
14	7	1	19.00	9.83
15	0	0	10.00	0.00
15	0	1	15.00	3.78
15	1	0	18.00	0.98
15	1	1	19.00	6.49
15	2	0	21.00	17.40
15	2	1	19.00	7.94
15	3	0	14.00	5.79

Acknowledgements

I take this opportunity of expressing my sincere thanks to Dr. Marjorie Harding and Dr. D.A. Rees for supervising my work during the past three years. I have appreciated their willingness to help and their careful guidance. I have also been grateful for the freedom they gave me to pursue those lines of research which seemed to be most interesting. Their friendliness and enthusiasm have always been encouraging.

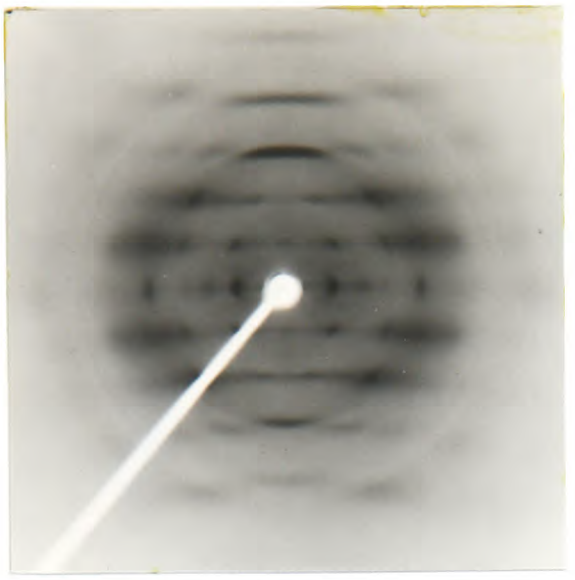
I would also like to give a special word of thanks to Dr. H.C. Watson of Bristol University for making available the diffractometer and I am grateful to him and to Dr. P. Wendell for their work in collecting and processing the data.

I would also like to thank Dr. W.E. Scott for his help, especially on the application of the symbolic addition procedure and Dr. N.S. Anderson, Dr. J.W.B. Samuel and Mrs. Carol McNab for preparing and photographing the polysaccharide fibres.

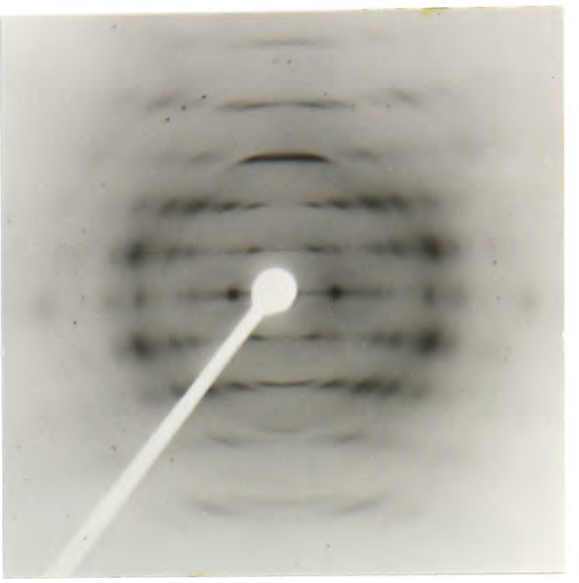
My thanks are also due to Dr. C.A. Beevers for the provision of molecular models and to him and my fellow crystallographers for their interest and their help.

I also wish to thank Dr. G.S. Pawley for the least squares refinement computer program and the Edinburgh Regional Computing Centre for computing facilities.

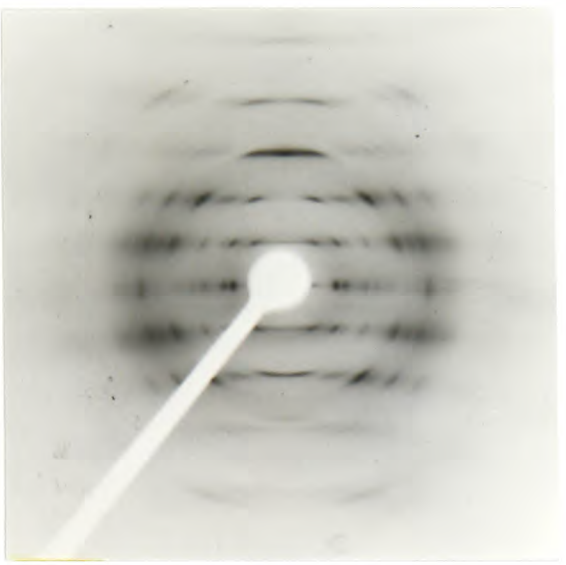
Finally I should like to thank Professor Sir Edmund Hirst and Professor E.A.V. Ebsworth for laboratory facilities, the Science Research Council for a Research Studentship which enabled me to do this work and Marine Colloids Inc. for an extension of the grant.



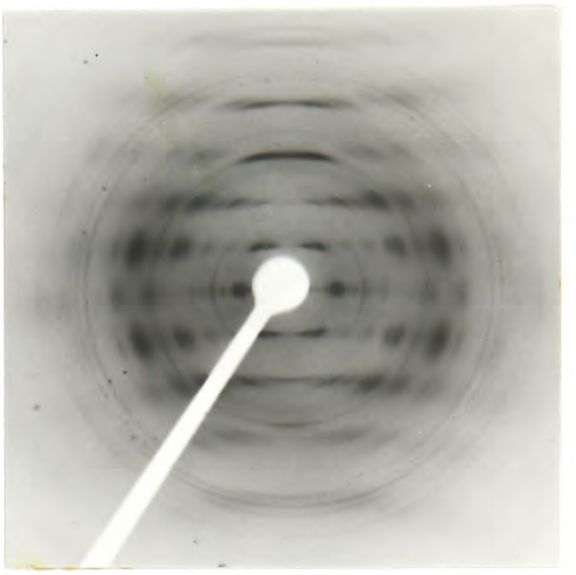
Mg^{++}



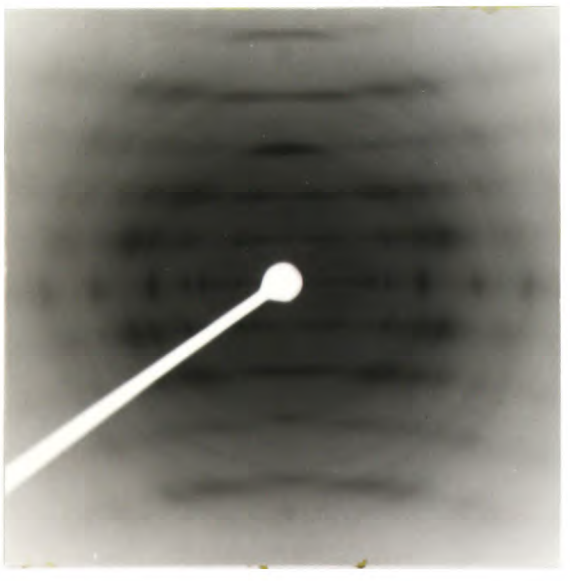
Li^{+}



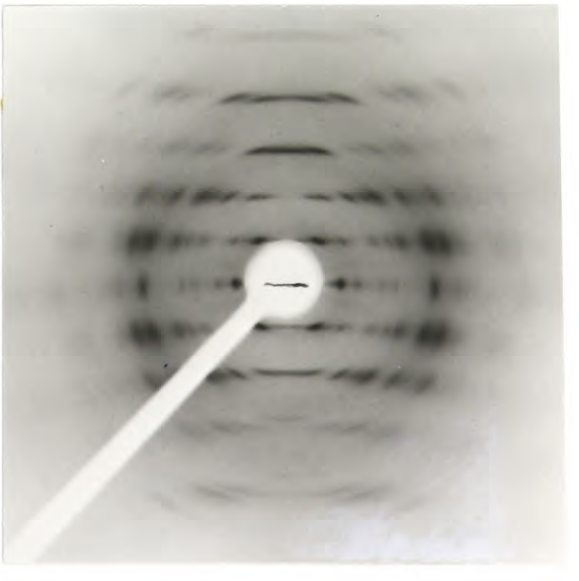
Na^{+} (I)



Na^{+} (II)

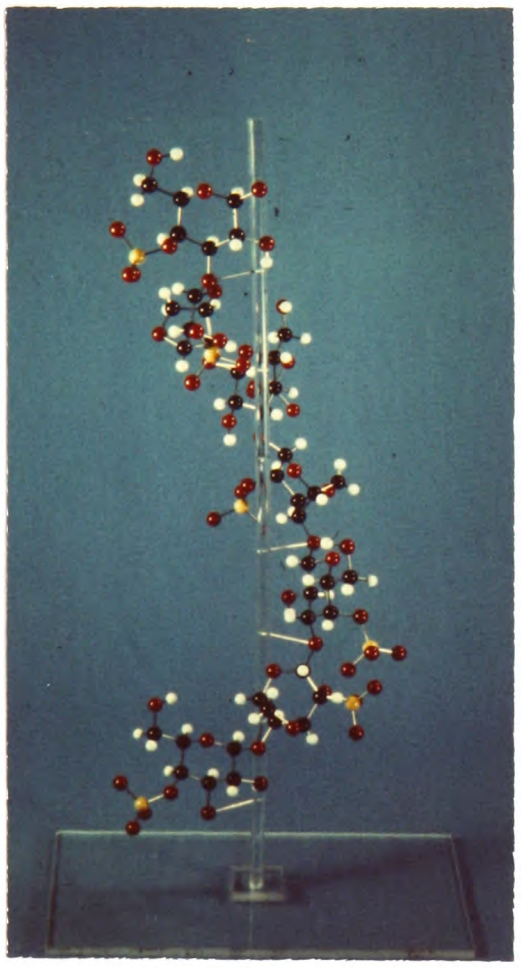


Sr^{++}

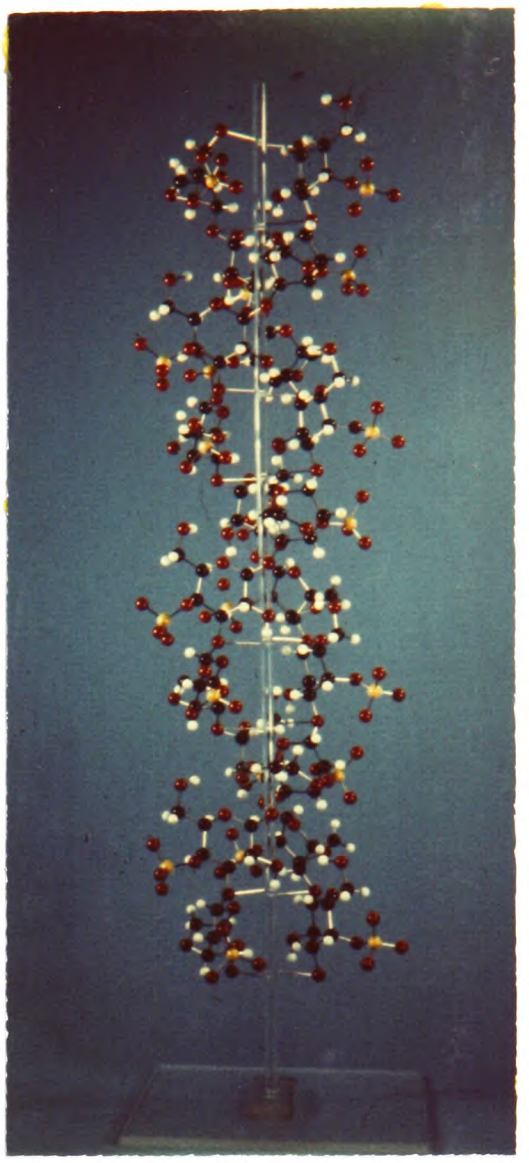


K^{+}

Fig. 49. X-ray diffraction photographs of iota fibres.



a.



b.

c.

Fig. 50. Models of iota:-
(a) A single strand of the double helix.
(b) & (c) The double helix.

Reprinted from *J. Mol. Biol.* (1969) **45**, 85–99

X-ray Diffraction Studies of Polysaccharide Sulphates: Double Helix Models for κ- and ι-Carrageenans

N. S. Anderson, J. W. Campbell, M. M. Harding,
D. A. Rees and J. W. B. Samuel

APPENDIX

Model Building Calculations for Alternating Polysaccharides

D. A. Rees

J. Mol. Biol. (1969) **45**, 85–99

X-ray Diffraction Studies of Polysaccharide Sulphates: Double Helix Models for κ- and ι-Carrageenans

N. S. ANDERSON, J. W. CAMPBELL, M. M. HARDING†,
D. A. REES† AND J. W. B. SAMUEL

Chemistry Department, University of Edinburgh, West Mains Road Edinburgh, EH9 3JJ, Scotland

(*Received 31 March 1969*)

Two polysaccharides have been studied which approximate in structure to alternating copolymers, $(—A—B)_n$, in which B is a residue of β-D-galactose-4-sulphate and A is a residue of 3,6-anhydro-α-D-galactose (κ-carrageenan) or its 2-sulphate (ι-carrageenan). The glycoside linkages are A1→3B and B1→4A. A small proportion of A residues also 'occur as α-D-galactose-6-sulphate and 2,6-disulphate in the natural polymers but were removed by fractionation or conversion to the anhydride with alkaline borohydride. X-ray diffraction photographs of oriented fibres of salts with various monovalent cations then showed fibre axis repeat distances of 24·6 Å (κ-carrageenan) and 13·0 Å (ι-carrageenan). Very similar models can be proposed for both polysaccharides, based on striking relationships between features in the various diffraction photographs, mathematical derivation of all models with appropriate dimensions and symmetry, and calculation of cylindrically averaged Fourier transforms. They are double helices with three disaccharide residues in a complete turn of each single chain, in 24·6 Å (κ-carrageenan) or 26·0 Å (ι-carrageenan). In ι-carrageenan the second chain is displaced exactly half a pitch from the first.

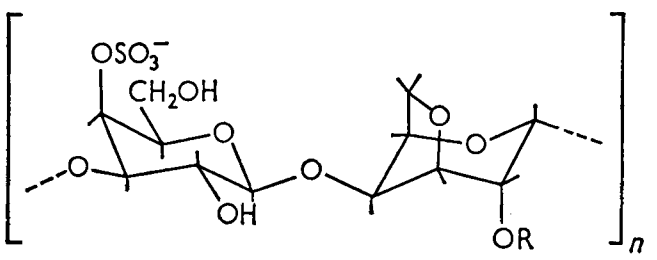
1. Introduction

Certain polysaccharides which we may call "gel-forming polysaccharides" exist in biological situations as swollen networks of hydrated chains, for example in animal and human cartilage, bacterial capsules, and the cell walls of young plants. Their function is partly structural but also involves their elasticity and water holding properties and ability to react selectively with cations. The "tie points" in such networks may be covalent in part, but secondary valence forces between chain segments are also likely to be important, and indeed offer mechanisms for subtle alteration of biological texture. It is therefore likely that a knowledge of polysaccharide tertiary structure, or inter-chain contacts, would be crucial to the understanding of the properties of these systems. Seaweeds provide many polysaccharides which have very striking properties as gels *in vitro* and which have primary structures regular enough to suggest the use of diffraction methods. We have selected two of this group as models in which to study the role of tertiary structure in polysaccharide networks.

Careful selection of source, sometimes followed by subfractionation and/or chemical modification of the native polysaccharide, is necessary to obtain regular structures likely to give interpretable fibre diagrams. Some years ago, before the polysaccharide structures were known, Bayley (1955) made an X-ray study. In so far as comparison

† All enquiries should be addressed to M.M.H. or D.A.R.

is possible, our results are in agreement but the interpretations are different. Our materials were prepared for X-ray diffraction from κ-carrageenan and ι-carrageenan as described in the Methods section, and henceforth referred to as "kappa" and "iota", respectively, although we did not use the crude native polysaccharides. As shown in formula I, these derivatives are regularly alternating polymers of the type $(—A—B)_n$ where B is the β-D-galactose-4-sulphate residue and A is the residue of 3,6-anhydro-α-D-galactose (in kappa) (Anderson, Dolan & Rees, 1968), or its 2-sulphate (in iota) (Anderson, Dolan & Rees, unpublished results).



I

Kappa: R=H

Iota: R=SO_3^-

2. Methods

(a) *Polysaccharide fibres*

ι-Carrageenan was prepared in the laboratory of Marine Colloids, Inc., from carefully handsorted *Eucheuma spinosum*. It was treated (Rees, 1961) with sodium hydroxide–sodium borohydride reagent at 80°C to convert the galactose-2,6-disulphate which represents about one tenth of the 4-linked residues (Anderson, Dolan & Rees, unpublished results), into 3,6-anhydride. Without this step the diffraction photographs of the potassium salt were almost featureless. Kappa was from *Chrondrus crispus* and was a highly purified subfraction described as sample 8 in an earlier paper (Anderson, Dolan, Penman *et al.*, 1968). The polysaccharides were obtained in each salt form by passage of a dilute solution through an ion exchange column in the appropriate state (Amberlite IR 120, analytical grade), followed by evaporation at 35°C under diminished pressure and final freeze drying. A drop of warm aqueous solution, at a concentration to give a soft gel when barely cool (1 to 4%), was placed across the glass beads in a cell (Fuller, Hutchison, Spencer & Wilkins, 1967) and drawn out whilst it set. Constant humidity (about 58%) was maintained by means of a saturated sodium bromide solution and the fibre was removed after 48 hr at 0 to 3°C. Fibre densities were measured by flotation in mixtures of carbon tetrachloride and benzene or carbon tetrachloride and methyl iodide.

(b) *X-ray diffraction*

Photographs were taken with CuKα radiation from a Philips fine focus X-ray tube, collimated by a fine lead glass capillary. The specimen was mounted on a Supper precession camera, with the fibre axis tilted, usually by 10°, out of the plane normal to the X-ray beam. The 'movement' of the precession camera was not used; the film was usually positioned 6·00 cm from the fibre. The intensity distributions were recorded with a Joyce–Loebl microdensitometer.

(c) *Transforms*

Intensity distributions were calculated on a KDF9 computer for various molecular models, according to the formula given by Klug, Crick & Wyckoff (1958) for the cylindrically averaged Fourier transform of a helical structure. Bessel functions up to the ninth order were included and the square of the Fourier transform was calculated at intervals of 0·02 in ξ.

3. Diffraction Photographs

(a) *Comparison of salt forms*

Various salts of kappa and iota gave fibres which were examined by X-ray diffraction:

Cation	Degree of orientation	
	Kappa salt	Iota salt
NH_4^+	+	+++
Li^+	none	+++
Na^+	none	+++
K^+	++	+++
Rb^+	++	++
Cs^+	+	+

The degree of orientation in the different salts, as judged by the X-ray photographs, varies with the ability to form a gel, and is generally better in the iota series than in the kappa series. The fibres could be described as semicrystalline (Holmes & Blow, 1965). The best fibres were those of the potassium salt of iota (Plate I).

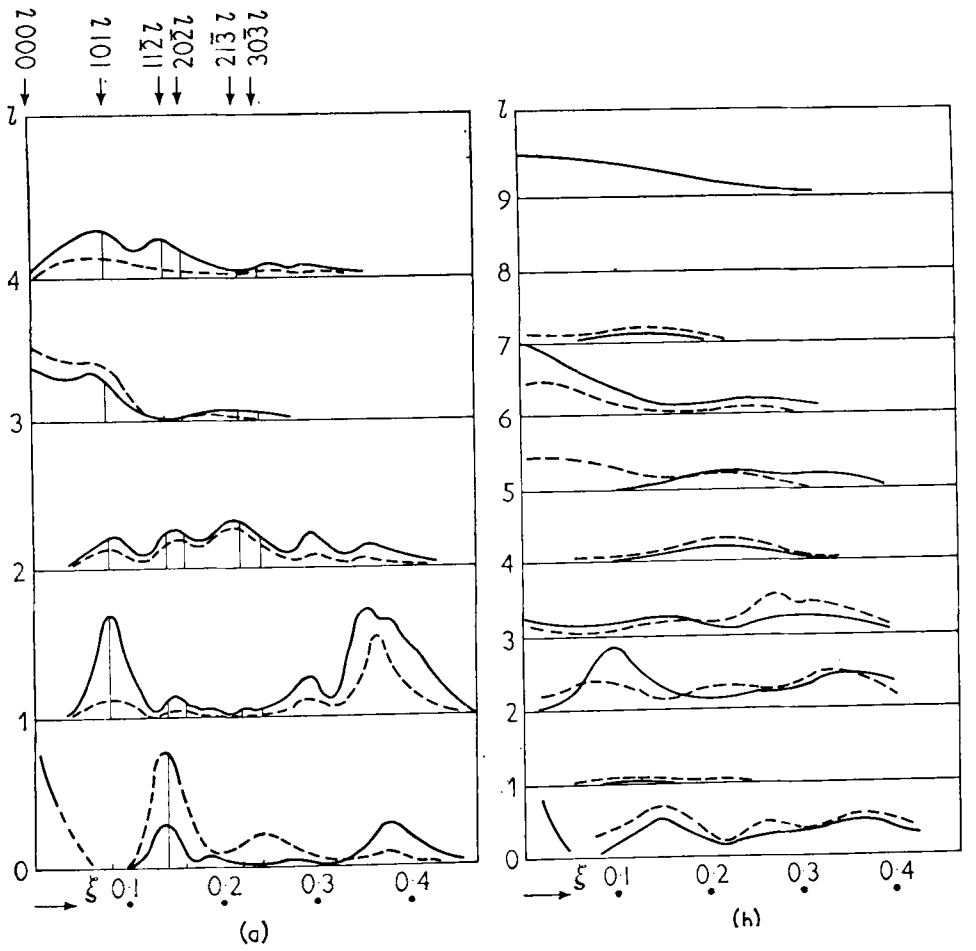


FIG. 1. Examples of intensity distribution in the layer lines (l) of kappa and iota salts. ξ is dimensionless and equal to 2 sin θ when $l = 0$. On the left are the ammonium and potassium salts of iota (broken lines and full lines), and on the right are the ammonium and rubidium salts of kappa (broken lines and full lines).

The positions of the first few reciprocal lattice points in iota corresponding to a hexagonal unit cell of side $a = 22{\cdot}6$ Å are indicated.

(b) *Fibre axis repeat distance for iota*

All the iota salts have a fibre axis repeat distance, c, of $13{\cdot}0 \pm 0{\cdot}3$ Å, deduced from layer line spacings. The diffraction pattern is consistent with a helical arrangement (Cochran, Crick & Vand, 1952), containing three disaccharide residues per turn of helix; on the meridian there is a diffraction maximum in the third layer but not in any other layer up to the fifth.

(c) *Salts are isostructural*

The fibre axis repeat distance is the same, within a few per cent, for the different salts of iota; the general character of the intensity distributions is the same but there are some marked differences, as shown in Figure 1. At present we attribute these to the different scattering powers of the cations and assume that the polysaccharide skeleton is the same throughout the iota salts or the kappa salts. Difficulties in deducing relative scale factors for the intensity distributions of the various salts have so far prevented us from using the formal isomorphous replacement method for structure solution.

(d) *Fibre axis repeat in kappa and relation of kappa to iota*

The diffraction patterns of the kappa salts are much more diffuse than those of the iota salts; the fibre axis repeat distance is $24{\cdot}6 \pm 0{\cdot}8$ Å, i.e. about twice that of iota. The intensity distributions indicate a close relation to iota; the layers $l = 0, 2, 4, 6$ of kappa are similar to the layers $l = 0, 1, 2, 3$ of iota while the layer $l = 1$ in all kappa salts is very weak or absent. Diffraction on the meridian—present only when $l = 3$, 6 or 9—shows that there are three disaccharide units per turn of this helix.

(e) *Unit cell*

The layer line spacings define the fibre axis repeat distance or the c axis of a unit cell. Within the layer lines the intensity distribution is nearly continuous, but in iota the first few maxima are consistent with a hexagonal cell of side $a = 22{\cdot}6$ Å—as shown in Figure 1. In the zero layer, $10\bar{1}0$ is absent whereas $11\bar{2}0$ is very strong; $11\bar{2}0$ corresponds to the first order of diffraction from the related hexagonal cell of side

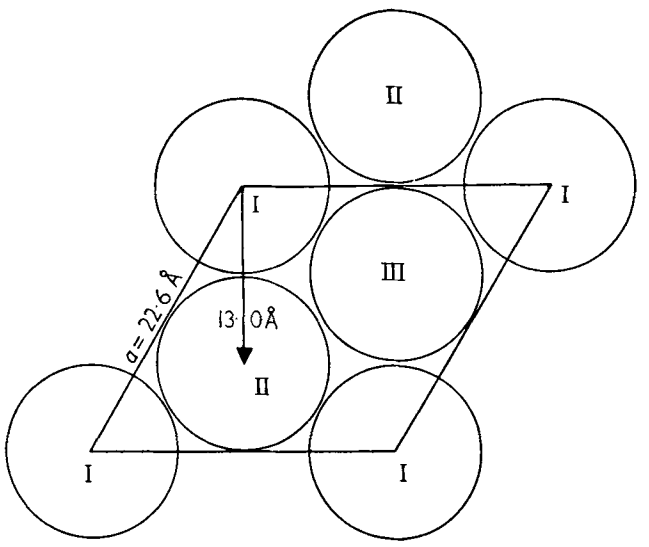


FIG. 2. Cross-section of a hexagonal array of helices of diameter about 13 Å. It is proposed that the helices I, II and III have different displacements in the c-axis direction so that, while the true unit cell has side $a = 22{\cdot}6$ Å, the unit cell seen in this projection has side 13·0 Å.

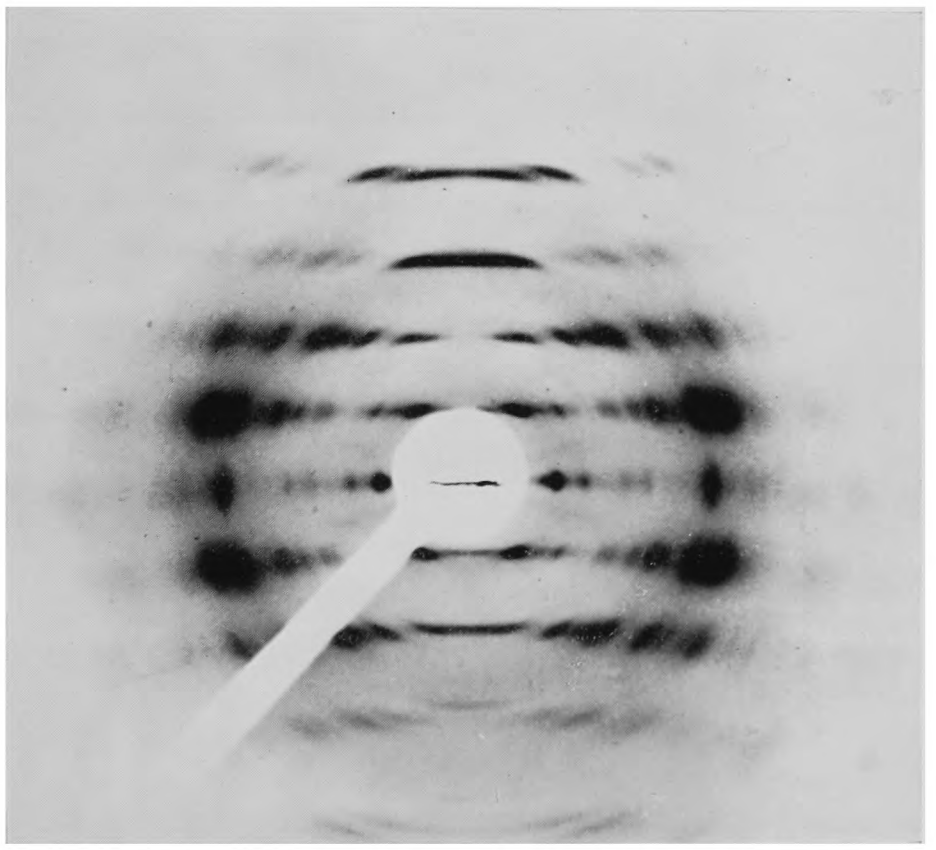


PLATE I. X-ray diffraction photograph of the potassium salt of iota carrageenan. Tilt of fibre was 10°.

TABLE 1

Numbers of disaccharide residues per fibre axis repeat from the measured densities

	Iota salts		Kappa salt
	NH_4^+	K^+	K^+
Molecular weight of disaccharide + 15% water	575	623	488
Observed density (g cm^{-3})	1·57	1·72	1·62
Hence volume per disaccharide ($Å^3$)	608	601	500
Unit cell a (Å)	22·6	22·6	*ca.*20
c (Å)	13·0	13·0	24·6
Unit cell volume/(3 × vol./disaccharide) = number of disaccharides per fibre axis repeat	3·15	3·19	5·68

$a/\sqrt{3}$=13·0 Å (Fig. 2) and suggests a hexagonal arrray of helices, each of diameter approximately 13 Å. Helices I, II and III in Figure 2 must have different displacements in the c direction. A similar cell with a approximately 20 Å is consistent with the kappa photographs.

The number of disaccharide units per fibre axis repeat in the above cells can be deduced from the measured densities, as shown in Table 1. Uncertainties in the water content and cell dimensions are sufficient to account for the discrepancy between the figures deduced and 3·0 or 6·0.

4. Double Helix Models

We propose that in both kappa and iota there are pairs of polysaccharide chains arranged as double helices. In each chemically distinct polysaccharide chain there are three disaccharides in one turn of helix, with a repeating distance of 26·0 Å (iota) or 24·6 Å (kappa). In iota, the disaccharide units of the second (parallel) helix are located exactly half way between those of the first helix giving a fibre axis repeat distance of 13·0 Å as shown in Figure 3. In kappa we do not yet know the exact relationship of the two helices except that it is not such as to halve the fibre axis repeat distance.

We first proposed this double helix model after consideration of the general features of the diffraction pattern and the close relation of kappa and iota, and then found that a model with acceptable stereochemistry in the polysaccharide chain could be built to conform to these requirements. For iota, a single helix with three disaccharide residues per turn, repeating after 13·0 Å, was still a formal possibility but should give a much larger intensity on the first layer line than that observed. A single helix structure for kappa, to maintain its similarity to iota, would need to have six disaccharide residues in *two* turns of a helix repeating after 24·6 Å; chemically identical groups would then be in stereochemically different situations. The double helix structure accounts very well for the weak first layer line in kappa.

Subsequently, we systematically explored polysaccharide chain conformations for iota which could give either single or double helices of the right dimensions and which had acceptable stereochemistry and van der Waals contacts (see below). The acceptable conformations could be represented by five single and six double helices, and

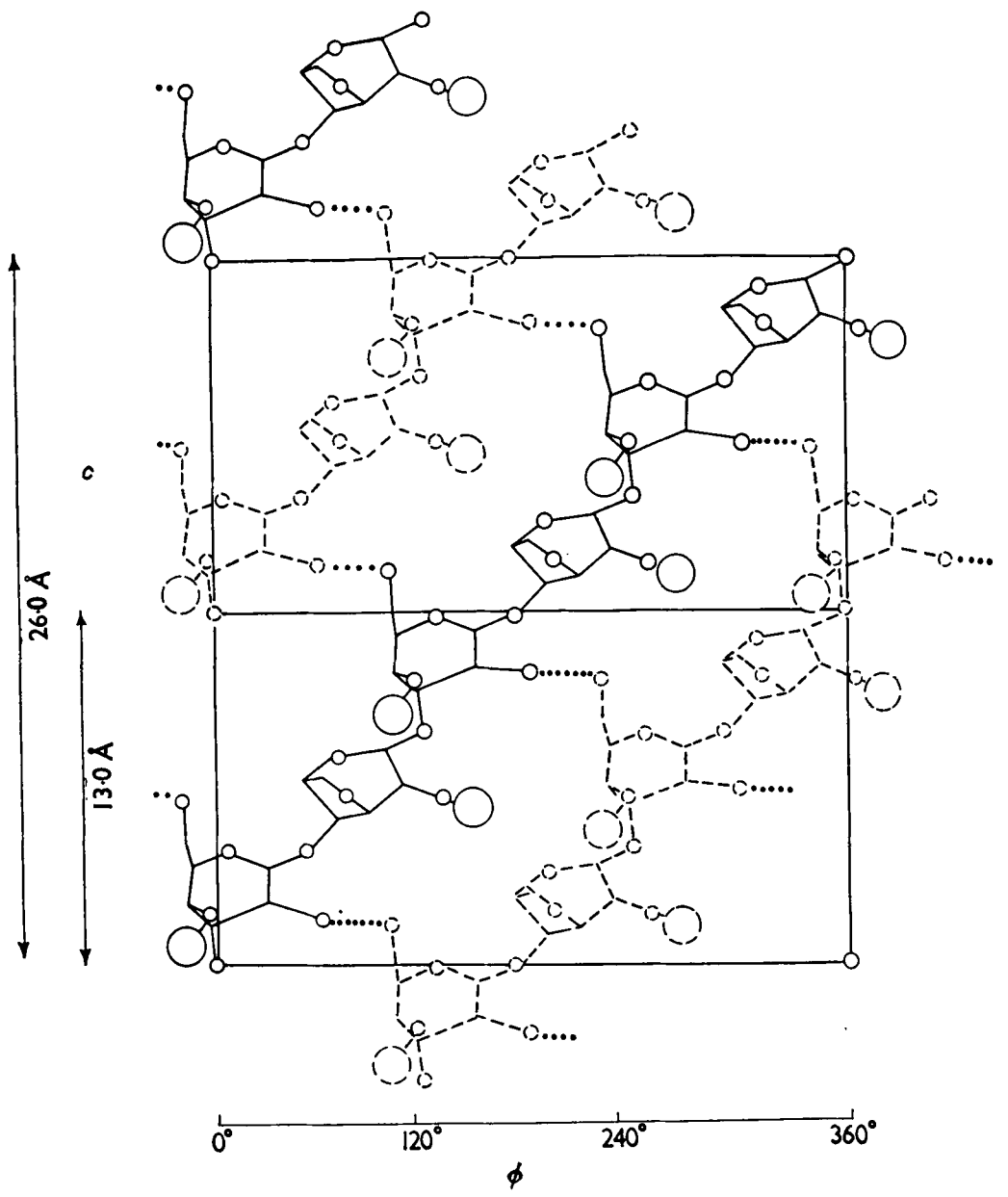


FIG. 3. The proposed arrangement of polysaccharide chains in iota. Two chemically distinct chains, represented by full and broken lines, respectively, and each having three disaccharide residues in one turn of helix of pitch 26·0 Å, are twisted together to form a double helix with a translation period of 13·0 Å. The structure is represented as a projection on the surface of a cylinder which has then been opened out. Small circles represent oxygen atoms and large circles are SO_3^- groups.

The co-ordinates correspond to model H (see text); the proposed O(2) .... O(6) hydrogen bond is shown by dotted lines.

sulphate ester group positions were calculated for each. Intensity distributions were calculated for all of these and compared with the Li^+ salt of iota. The agreement between observed and calculated intensity was poor for all the single helices, better for most of the double helices, and quite promising for models F, G and H (Table 2). Two of the intensity distributions are illustrated in Figure 4. The best model appeared to be H, which was essentially the same as the model, I, derived by a different mathematical procedure and selected as the best according to stereochemical criteria. Inspection of a molecular model built from the calculated co-ordinates suggested a hydrogen bond, perpendicular to the helix axis, between O(6) of one galactose unit and O(2) of its partner in the second strand. The only hydroxy-groups in the repeating

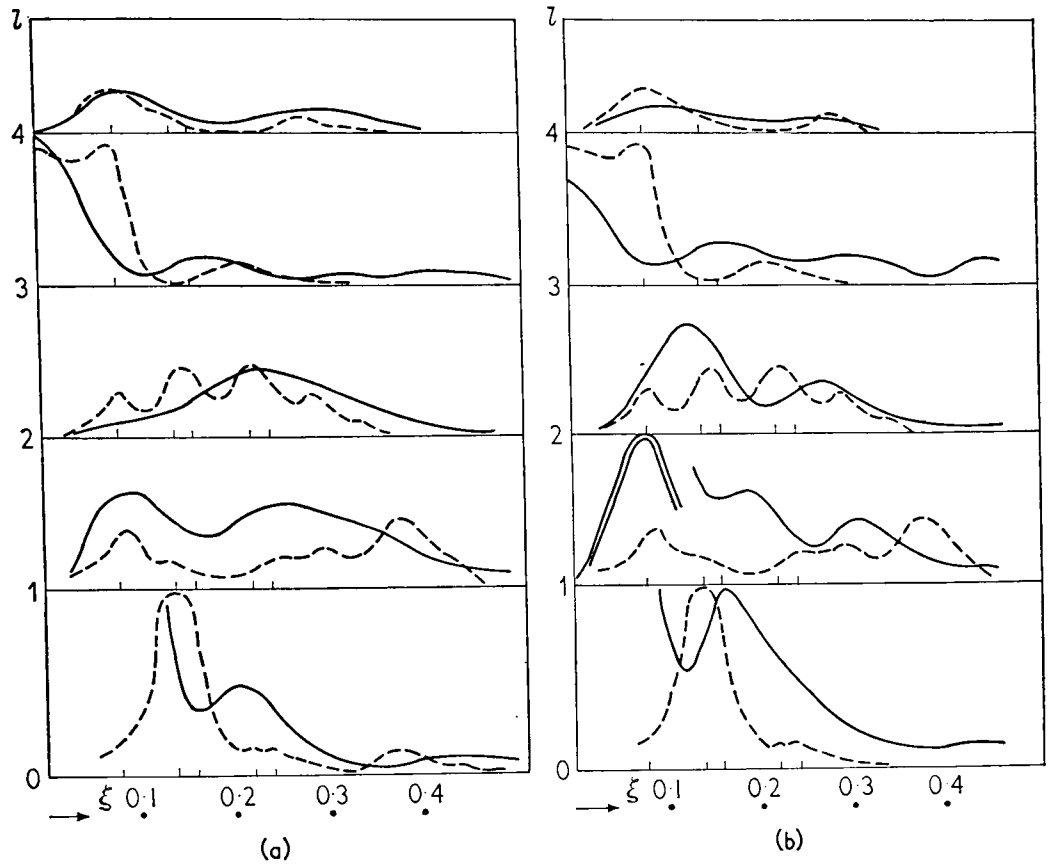


FIG. 4. Comparison of observed intensity distribution (broken lines) in the lithium salt of iota with the intensity distribution calculated for (a) the double helix model H, (b) the single helix model A. The latter model is rejected, mainly on account of the first layer disagreement. (The positions of reciprocal lattice points are indicated as in Fig. 2.)

unit (see formula I) would then be engaged in hydrogen bonding, making this helix seem very plausible on structural grounds. There is some experimental evidence for such a hydrogen bond (D. A. Rees & F. B. Williamson, unpublished results). Oriented films of kappa and iota show a continuous and almost featureless absorption in the O—H stretching region of the infrared spectrum but, after deuteration, there is a sharp band at 3330 cm^{-1} which corresponds to a hydrogen bond of length 2·8 to 2·9 Å. The very pronounced perpendicular dichroism of the band and its resistance to deuteration are consistent with an inter-strand hydrogen bond buried in the helix interior. Its presence in the spectra of kappa and iota is a further indication of similarity between the two conformations.

We therefore conclude that the structure is a double helix with a conformation close to H; this is illustrated in Figure 3, and the model I, which is almost identical, is illustrated in Figure 5. The sulphate groups are on the outside of the double helix and cannot be as rigidly fixed by stereochemical requirements as can most other atoms of the polysaccharide chain. However, the average contribution of the sulphate groups to the calculated intensity distribution is about half the total. This distribution is therefore very sensitive to the sulphate positions and not as sensitive as we would wish to the conformation of the polysaccharide chain. We expect that the hexagonal array of double helices is held together by interactions between the sulphate groups and cations. We are now trying to establish these positions and hope that this will lead to better agreement between the observed and calculated intensity distributions.

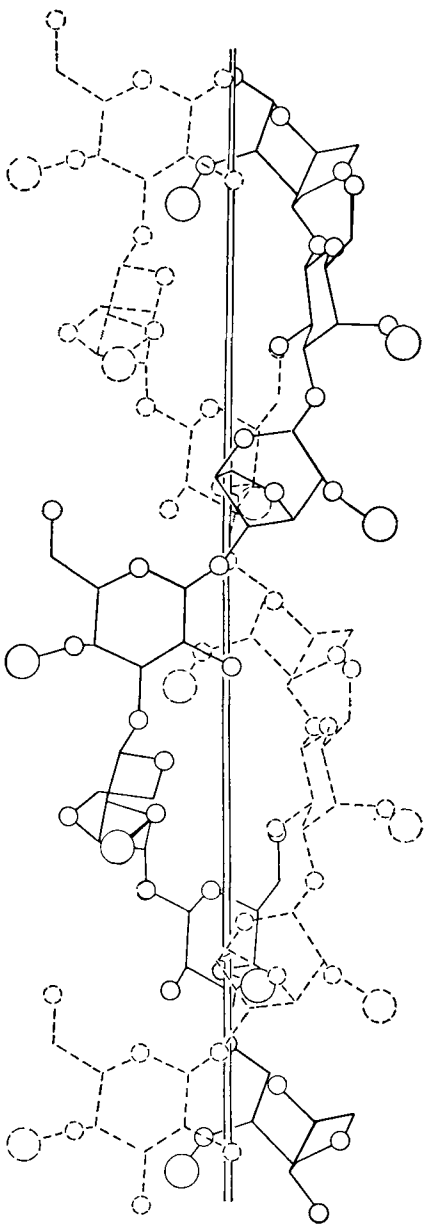


FIG. 5. Model I for the iota double helix. Full lines and dotted lines represent chemically distinct chains. Small open circles represent oxygen atoms and larger circles SO_3^- groups. See Fig. 3 for the position of O(6) .... O(2) hydrogen bonds.

5. Model Building

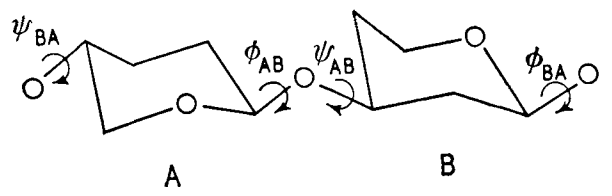
The feasibility of a double helix was confirmed for iota with ball-and-spoke and with Corey–Pauling–Koltun space filling models. Two independent computer methods were then used in an objective and exhaustive survey, each based on a different mathematical treatment and on a separate set of reasonable stereochemical assumptions and residue co-ordinates.

It is assumed that each monosaccharide residue, and the bond angle at each glycosidic oxygen, has a fixed geometry which can be predicted from analogous crystal structures. To define a regular chain conformation it is then only necessary to fix the four angles of rotation, ϕ_{AB}, ψ_{AB}, ϕ_{BA} and ψ_{BA}, which are shown in formula II†. This approach has been much used in conformational analysis of macromolecules

† For consistency with the literature, we must unfortunately use ϕ in three different ways: (i) ϕ_{AB}, ϕ_{BA} are torsion angles about glycosidic bonds; (ii) ϕ_i are angles between vectors and bonds (see Appendix); (iii) ϕ is a cylindrical polar co-ordinate.

(Ramachandran, Ramakrishnan & Sasisekharan, 1963; De Santis, Giglio, Liquori & Ripamonti, 1963; Ramachandran & Sasisekharan, 1968; Scheraga, 1968; Brant & Flory, 1965).

The first computer method, "Modelbuild I", was used only to derive a sterically feasible iota double helix; the mathematics will be described elsewhere. All possible



II

combinations of values of ϕ_{AB} and ψ_{AB} were explored and those which led to impossible contacts across the glycosidic linkage, judged by "outer limit" criteria (Ramachandran *et al.*, 1963), were rejected. For each allowed combination of these angles, the values from the diffraction photograph for n (screw symmetry) and h (projected length of the disaccharide residue on the helix axis) were used to fix the remaining parameters, ϕ_{BA} and ψ_{BA}. The conformation was accepted if contacts were satisfactory (i) across the B—A linkage (ii) with a second chain added to form a double helix and displaced from the first by half the translation period. Some of the principles involved have been used before in constructing trial models for cellulose (Jones, 1958), xylan (Settineri & Marchessault, 1965), and amylose triacetate (Sarko & Marchessault, 1967), although each of these structures has only one monosaccharide residue in the chemical repeating unit and therefore the conformation has only two variable torsion angles.

Most of the carrageenan structures were derived by the second procedure, Modelbuild II. The stereochemical assumptions, and conformational maps which show allowed combinations of each ϕ and ψ, have been described (Rees, 1969*a*). These combinations were taken together in all possible ways to compute the helical parameters, n and h, and thus, by interpolation, n and h values to match experiment. The coordinates of the polysaccharide chain were then transferred to helix axes for further calculations. Details are given in the Appendix. All interatomic distances were checked to the fourth residue up and down the chain. Contacts between the two chains in the double helix as proposed for iota were checked; in kappa the screwing of one chain relative to the other must be considered and this was explored with the typical results shown in Figure 6. Sulphate esters have some rotational freedom but Corey–Pauling–Koltun models suggest that this is confined to an oscillation through $\pm 10°$ about a mean position in which the H — C — O — S system is coplanar and the C — H and O — S bonds are eclipsed. The sulphates were placed in this conformation with one of the remaining O — S bonds also in the plane and *anti* to the C — O bond. The bond angles were tetrahedral with S — O bond lengths of 1·60 (to the ester oxygen) and 1·45 Å (to the terminal oxygens).

In the absence of any constraints whatsoever and if the angles were taken in 10° steps, the total number of regular conformations for a carrageenan chain would be about 1,680,000. When the outer limit criteria were imposed with the requirement that $n = \pm 3$ and $h = 8{\cdot}67$ Å, Modelbuild I reduced this number to 25 right-handed and 15 left-handed helices. All the left-handed helices failed to accommodate a second

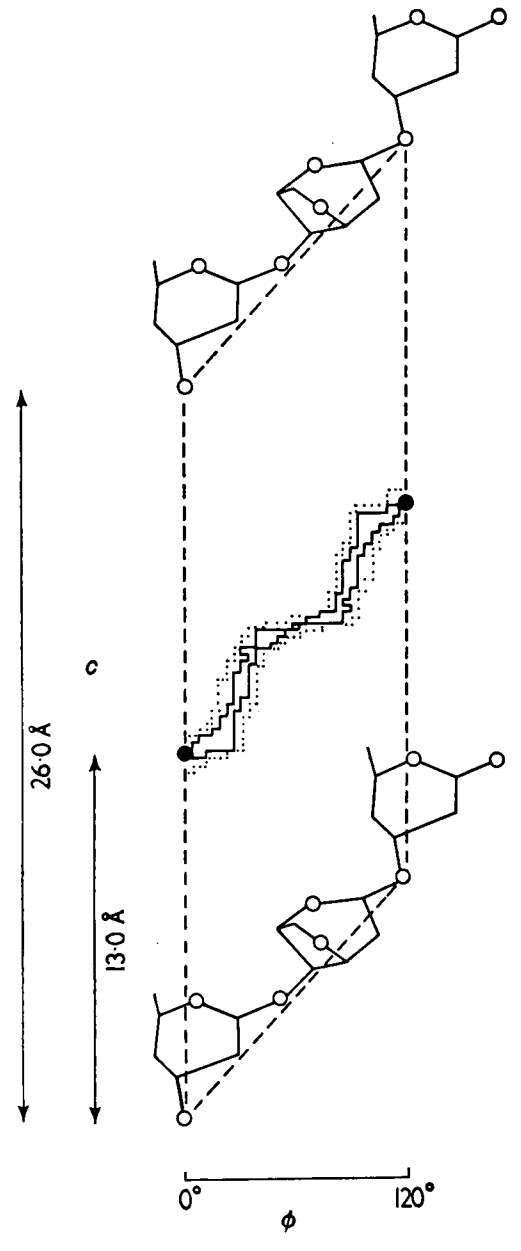


FIG. 6. Computer calculations to show that it is sterically possible for carrageenans to form double helices with the chains in different relative positions. An iota chain is drawn in radial projection (cf. Fig. 3) to correspond to model H with some atoms omitted for clarity. A disaccharide residue, representing a parallel second strand, is moved over the grid inside the large parallelogram defined by the dashed lines, with checking of all relevant contacts to the first chain. Sulphate groups were omitted at this stage; they are on the outside of the double helix and Corey–Pauling–Koltun models show that they have little steric influence on packing within the double helix. Possible positions for the origin atom (O(3) of B) occur in a strip across the middle of the parallelogram and are boxed by solid lines ("fully allowed") and dotted lines ("marginally allowed"). Contact criteria were as used by Rees (1969*a*).

The double helix proposed for iota has the origin of the second chain in the fully allowed position marked thus ●. Allowed areas of about the same size, but differing in shape and position, were also found in the middle of the map when the disaccharide was antiparallel to the first chain.

chain to form a double helix but 17 right-handed helices remained. When calculations were repeated with the allowable contact distances increased by 0·1 Å, all the conformations failed except one, designated "model I".

Using Modelbuild II, systematic variation of ϕ_{AB}, ψ_{AB}, ϕ_{BA} and ψ_{BA} in 20° steps followed by interpolation and checking of remote contacts and contacts between strands, led to 17 iota double helices, all of which had right-handed chains. Comparison

of conformational maps calculated at 10° and 20° intervals showed that no allowed areas were missed by using the larger steps. Conformations which differed by less than 15° in all angles of rotation were grouped and one representative was taken from near the middle of the zone to give six final possibilities (Table 2); model H resembled

TABLE 2

Representative iota conformations derived by systematic exploration of all possibilities using Modelbuild II

Model	Type	Helical parameters		Conformation angles (in degrees)			
				A — B linkage		B — A linkage	
		n	h	ϕ_{AB}	ψ_{AB}	ϕ_{BA}	ψ_{BA}
A	Single	3·00	4·33	110·0	221·0	97·6	142·0
B	Single	3·00	4·33	120·0	218·0	74·5	161·0
C	Single	3·00	4·33	120·0	225·9	103·0	120·7
D	Single	3·00	4·33	101·0	229·5	120·0	123·3
E	Single	−3·00	4·33	123·5	241·0	305·1	180·0
F	Double	3·00	8·67	130·0	191·12	142·75	160·0
G	Double	3·00	8·67	122·3	194·0	160·0	151·3
H	Double	3·00	8·67	139·7	198·2	148·0	140·0
J	Double	3·00	8·67	159·9	204·2	140·0	122·0
K	Double	3·00	8·67	160·0	190·0	119·6	153·4
L	Double	3·00	8·67	171·0	195·8	120·0	137·7

A negative value of n implies a left-handed helix; all other helices are right handed.

model I. The average radii of all these double helices were similar and would allow packing in a unit cell of the type proposed, with space for cations and water; most atoms were within 5 Å of the helix axis with only a few as far out as 7 Å, but no further.

A similar number of kappa conformations could be derived, taking $h = 8{\cdot}20$ Å. Kappa and iota possibilities usually differed by bond rotations of less than 15°, and were therefore similar in their chain contours. Left-handed kappa helices, like the corresponding iota helices, were excluded by their failure to accommodate a second chain even when outer limit criteria were used in the test.

Single helices were also considered for iota, taking $n = \pm 3$ and $h = 4{\cdot}33$; 15 conformations were found initially but seven of these were eliminated because of remote contacts closer than outer limit distances and others were similar enough to be grouped, giving the five final possibilities listed in Table 2, only one of which was left-handed. They had larger radii than the double helices, with many atoms placed 5·5 to 9 Å from the axis, and would be more difficult to pack in a unit cell with the dimensions proposed.

In the maps for inter-chain contacts in various double helices derived for kappa and iota, the chains seemed to be least crowded when symmetry related as proposed for iota (see Fig. 6). In kappa, the chains may be parallel and displaced from the iota arrangement by "screwing", or they might be antiparallel. Enough maps have been calculated to show that adequate space for the polysaccharide chain is available in either case over a range of screw displacements, but further progress is unlikely until the sulphate and cation positions have been deduced.

6. Possible Relation to Physical and Biological Properties

These double helices are the first examples for polysaccharides, although a triple helix has been proposed for an algal xylan (Atkins, Parker & Preston, 1969).

Both κ- and ι-carrageenans occur naturally in the gel state and, after extraction, may form stiff thermally reversible gels in low concentrations in aqueous solution (Percival & McDowell, 1967; Rees, 1969*b*). It has been shown that this gelation involves the cross-linking or close association of two or more chains (Anderson, Dolan, Penman *et al.*, 1968). The similarity between the cation requirements for gelation and fibre orientation by κ-carrageenan suggests the same type of association in each situation. Gelation is not prevented by formal removal of a proportion of either type of sulphate shown in formula I (Painter, 1960; Stancioff & Stanley, 1969) but formal transfer of sulphate from C(4) (as in I) to C(2) of B, or formal insertion of a proportion of galactose-6-sulphate in place of A residues, is inhibitory (Anderson, Dolan, Lawson, Penman & Rees, 1968; Lawson & Rees, 1968; Anderson, Dolan, Penman *et al.*, 1968). All these effects are understandable in terms of the double helix model and we therefore propose that gel melting and setting proceed by processes shown in Figure 7,

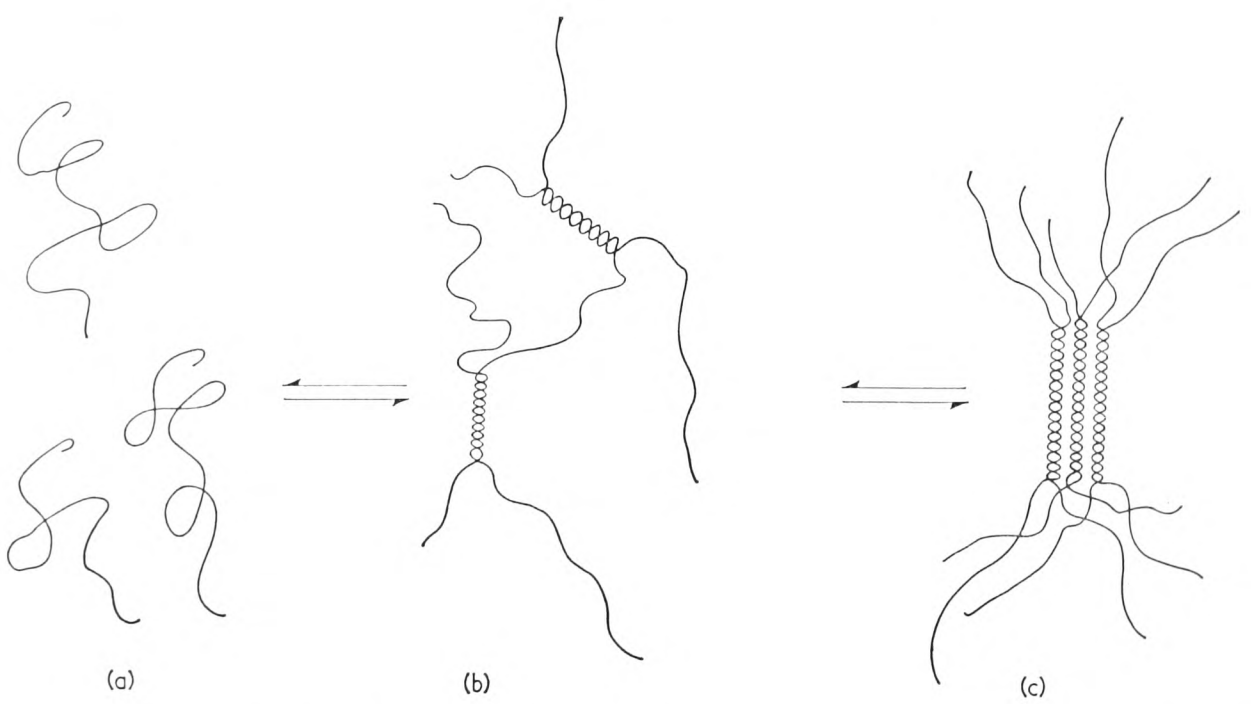


FIG. 7. Suggested relation of polysaccharide chains (a) in solution, (b) and (c) in the gel. The extent to which the helices are aggregated in the gel is not yet known, but (c) must represent the oriented fibre.

and that biological texture is regulated by control of helix content *via* primary structure. Other polysaccharide sulphate networks, of animal and seaweed origin, have primary structures based on an alternating arrangement of 1,3 and 1,4 linkages. Model building suggests that their ranges of allowed conformations are similar to the carrageenans (Rees, 1969*a*) and possible relationships are now under experimental investigation.

We are grateful to Dr S. Arnott for advice and helpful discussions, to Mrs C. S. Banks for measuring fibre densities, to the Science Research Council for studentships to two of us

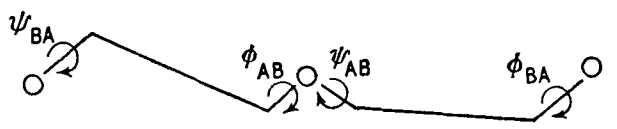
(J. W. C. and J. W. B. S.), to Marine Colloids, Inc., for a studentship to one of us (N. S. A.) and for polysaccharide samples, to the Edinburgh Regional Computing Centre for their co-operation, and to Professor Sir Edmund Hirst for his interest and encouragement.

APPENDIX

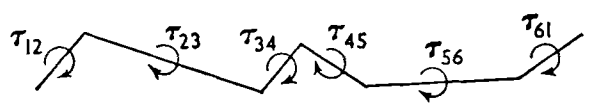
Model Building Calculations for Alternating Polysaccharides

D. A. Rees

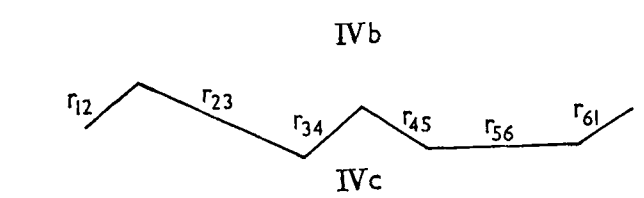
The following procedure ("Modelbuild II") can be applied to any alternating polysaccharide chain without 1,5 and 1,6 linkages, in any conformation. It is illustrated here by reference to carrageenans. The chain (II) is treated as a backbone (III) on which other groups are fixed appendages. Following the methods and notations of Sugeta & Miyazawa (1967), the helical parameters, n and h, can be expressed in terms of the angles of internal rotation τ_{ij} (see IVa), the angles between vectors and bonds ϕ_i (see IVb), and the interatomic distances r_{ij} (see IVc). All these quantities are implicit in the starting assumptions (Rees 1969*a*; and also p. 92) except τ_{12}, τ_{34}, τ_{45} and τ_{61} which correspond, respectively, to ψ_{BA}, ϕ_{AB}, ψ_{AB} and ϕ_{BA}. A rotation matrix, **T**, can also be derived for use in transferring the polysaccharide in any conformation to the helix axes defined by Sugeta & Miyazawa (1967).



III



IVa



IVb

IVc

The calculation proceeds in the stages described below; stages 1 to 6 were done together in one computer programme, stage 7 by hand, and stages 8 to 13 in a second programme.

(1) Co-ordinates of sugar residues A and B were obtained on axes placed at each O(1), as defined before (Rees, 1969*a*).

(2) A second set of co-ordinates was generated, for each residue, referred to axes at each non-reducing glycosidic oxygen with Ox along the corresponding O—C bond (O(3)—C(3) for B and O(4)—C(4) for A); each xy plane is defined by C(1) so that this atom has a positive y co-ordinate.

(3) The fixed parameters, τ_{ij}, ϕ_i, r_{ij}, were calculated or (for ϕ_1 and ϕ_4) assigned.

(4) Using the initial co-ordinates for A and the co-ordinates of B from stage 2, the combinations of τ_{34} and τ_{45} were obtained which are not excluded by outer limit criteria. Each angle was varied stepwise from 0° to 360° for every value of the other and the B co-ordinates were referred to the A axes at each stage, by successive rotation of the B axes about Ox through τ_{45}, about Oz through $\pi - \phi_4$ and then about Ox through τ_{34}; this allows all interatomic distances to be calculated so that sterically impossible conformations can be rejected.

(5) A list of all possible combinations of τ_{12} and τ_{61} was obtained similarly.

(6) Each allowed combination of τ_{34} and τ_{45} was taken with each allowed combination of τ_{12} and τ_{61}, and with the fixed values from stage 3, to calculate n and h.

(7) The results from stage 6 were plotted on $n - h$ graphs (Ramachandran *et al.*, 1963), holding two angles constant at a time; n and h were found to change smoothly with the torsion angles. The conformations required were at the intersections of the $n = \pm 3$ curves with the curve corresponding to the value of h derived from the diffraction photograph. Some recalculation around the intersection was usually required to locate these conformations accurately. The process was repeated for all six pairs of angles. The result is a list of sets of τ_{12}, τ_{34}, τ_{45}, and τ_{61} values which would give the measured or postulated n and h.

(8) For each τ_{ij} set from the previous stage, a disaccharide residue (B — A) in the corresponding linkage conformation (τ_{61}, τ_{12}) was generated on the B axes of stage 2. This residue was then rotated about Ox, in the appropriate sense, through τ_{45}. The result is the chemical repeating unit of the polysaccharide with respect to fixed molecular axes and in a conformation which corresponds to a particular τ_{ij} set.

(9) Each τ_{ij} set was used with the parameters from stage 3 to calculate the elements of the **T** matrix (Sugeta & Miyazawa, 1967) which rotates the molecular axes to make them parallel with the helix axes, and ρ, the distance between the origin of the molecular axes and the z axis of the helix.

(10) The results of the preceding stage were used to transfer the disaccharide repeating unit (stage 8) to the helix axes and hence, with n and h, to generate a complete turn of helix, or more.

The checking of contacts between non-adjacent residues and residues in different chains, and the placing of substituents such as sulphate esters, then follows as outlined in the main part of the paper.

REFERENCES

Anderson, N. S., Dolan, T. C. S., Lawson, C. J., Penman, A. & Rees, D. A. (1968). *Carbohydrate Res.* **7**, 468.

Anderson, N. S., Dolan, T. C. S., Penman, A., Rees, D. A., Mueller, G. P., Stancioff, D. J. & Stanley, N. F. (1968). *J. Chem. Soc.* (*C*), 602.

Anderson, N. S., Dolan, T. C. S. & Rees, D. A. (1968). *J. Chem. Soc.* (*C*), 596.

Atkins, E. D. T., Parker, K. D. & Preston, R. D. (1969). *Proc. Roy. Soc.* B **173**, 209.

Bayley, S. T. (1955). *Biochim. biophys. Acta,* **17**, 194.

Brant, D. A. & Flory, P. J. (1965). *J. Amer. Chem. Soc.* **87**, 2788.

Cochran, W., Crick, F. H. C. & Vand, V. (1952). *Acta Cryst.* **5**, 581.

De Santis, P., Giglio, E., Liquori, A. M. & Ripamonti, A. (1963). *J. Polymer. Sci.* Pt. A, **1**, 1383.

Fuller, W., Hutchison, F., Spencer, M. & Wilkins, M. H. F. (1967). *J. Mol. Biol.* **27**, 507.

Holmes, K. C. & Blow, D. M. (1965). *Methods Biochem. Analysis,* **13**, 113.

Jones, D. W. (1958). *J. Polymer Sci.* **32**, 371.

Klug, A., Crick, F. H. C. & Wyckoff, H. W. (1958). *Acta Cryst.* **11**, 199.
Lawson, C. J. & Rees, D. A. (1968). *J. Chem. Soc.* (*C*), 1301.
Painter, T. J. (1960). *Canad. J. Chem.* **38**, 112.
Percival, E. E. & McDowell, R. H. (1967). *Chemistry and Enzymology of Marine Algal Polysaccharides.* London: Academic Press.
Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). In *Aspects of Protein Structure,* ed. by G. N. Ramachandran, p. 121. New York: Academic Press.
Ramachandran, G. N. & Sasisekharan, V. (1968). *Advanc. Protein Chem.* **23**, 283.
Rees, D. A. (1961). *J. Chem. Soc.* 5168.
Rees, D. A. (1969*a*). *J. Chem. Soc.* (*B*), 217.
Rees, D. A. (1969*b*). *Advanc. Carbohyd. Chem.* **24**, in the press.
Sarko, A. & Marchessault, R. H. (1967). *J. Amer. Chem. Soc.* **89**, 6454.
Scheraga, H. A. (1968). *Advanc. Phys. Org. Chem.* **6**, 103.
Settineri, W. J. & Marchessault, R. H. (1965). *J. Polymer. Sci.,* Pt. C, **11**, 253.
Stancioff, D. J. & Stanley, N. F. (1969). *Internat. Seaweed Symp.* vol. 6, in the press. Oxford: Pergamon Press.
Sugeta, H. & Miyazawa, T. (1967). *Biopolymers,* **5**, 673.